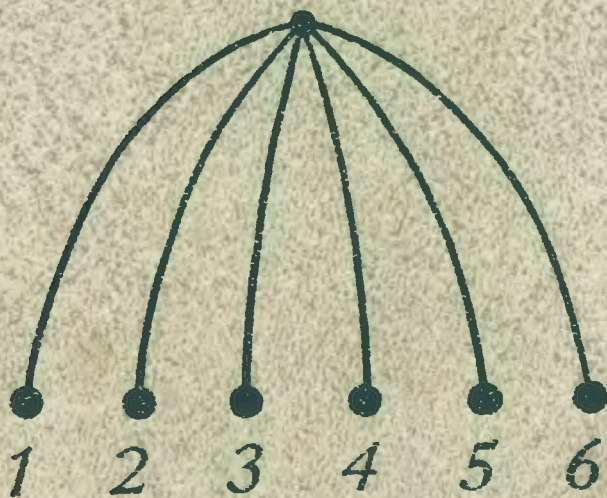


М. Кимура

*Молекулярная
эволюция:
теория
нейтральности*



The Neutral Theory of Molecular Evolution

Motoo Kimura

National Institute of Genetics, Japan

Cambridge University Press
Cambridge
London New York New Rochelle
Melbourne Sydney

М. Кимура

Молекулярная
эволюция:
теория
нейтральности

Перевод с английского
канд. биол. наук А. Г. Имашевой,
канд. биол. наук Е. А. Салменковой
и
Е. Я. Тетушкина

под редакцией
д-ра биол. наук Ю. П. Алтухова
и
д-ра биол. наук Л. А. Животовского

УДК 575.17+575.8
К40
ББК 28.04

Кимура М.

К40 Молекулярная эволюция: теория нейтральности: Пер. с англ. — М.: Мир, 1985. — 394 с., ил.

Монография японского ученого посвящена созданной им теории эволюции на молекулярном уровне. В книге рассмотрены история эволюционного учения, роль нейтральных мутаций в эволюции, молекулярные аспекты популяционной генетики. Предназначена для эволюционистов, генетиков, молекулярных биологов.

К $\frac{2001000000-429}{041(01)-85}$ 123—85, ч. 1.

ББК 28.04
57.023

Редакция биологической литературы

Предисловие редакторов перевода

Эта книга написана одним из ведущих специалистов в области теоретической популяционной генетики, много сделавшим для развития целого ряда ее направлений. Особенно известны работы М. Кимуры по созданию так называемых «диффузионных моделей», описывающих стохастические процессы эволюции генетической структуры популяций ограниченной численности. Фактически он создал новое направление в науке, применив принципы популяционной генетики к новейшим достижениям генетики молекулярной, прежде всего таким, как обнаружение необычайно широкого полиморфизма белков, расшифровка первичной структуры белков и нуклеиновых кислот, выявление экзонно-интронной организации генов. Опираясь на экспериментальные данные молекулярной биологии и разработанный им математический аппарат, М. Кимура обосновал оригинальную концепцию, широко известную как «теория нейтральности молекулярной эволюции».

Согласно представлениям автора, в основе эволюционных изменений на уровне белков и нуклеиновых кислот лежит фиксация нейтральных или почти нейтральных (по отношению к отбору) мутаций в результате случайного дрейфа генов. Эта идея, впервые сформулированная им в 1968 г., сразу же стимулировала широкие исследования полиморфизма белков и их эволюции, вызвала не прекращающуюся и по сей день плодотворную дискуссию, оказав тем самым существенное влияние не только на генетику популяций, но и на эволюционную биологию в целом.

Важно отметить, что теория нейтральности построена на строгой математической основе и, следовательно, позволяет проводить количественное изучение эволюции генов и белков и сопоставлять теоретические выводы с фактическими данными.

В книге детально разобраны все доводы «за» и «против» теории нейтральности и достаточно убедительно показано, что она объясняет многочисленные факты, полученные в различных областях молекулярной биологии и генетики. Книга написана логически очень четко, что позволяет ясно представить себе основную мысль автора и те ключевые факты, которые легли в основу его концепции.

Развиваемые в данной книге взгляды иногда рассматриваются как антитеза дарвинизму, что, конечно, не соответствует

действительности. Автор сам неоднократно подчеркивает, что его теория относится только к механизмам *молекулярной* эволюции на уровне нуклеотидных и аминокислотных замен, ни в коей мере не подменяя теории прогрессивной эволюции, созданной главным образом Ч. Дарвином. Из книги следует, что теория нейтральности возникла не на голом месте, и М. Кимура совершенно справедливо отдает дань пионерским работам Р. Фишера, Дж. Холдейна, С. Райта, Г. Мёллера, Ф. Добржанского и др. Отметим, однако, что он совершенно не упоминает имен С. С. Четверикова, И. И. Шмальгаузена, А. С. Серебровского, а также их учеников и последователей, заложивших основы современного учения о популяциях и факторах их генетических изменений.

Несколько слов об авторе. Мотоо Кимура возглавляет отдел популяционной генетики в Национальном институте генетики в Мисиме. Он написал ряд книг, в том числе «Диффузионные модели в популяционной генетике» (1964), «Введение в теорию популяционной генетики» (1970), «Теоретические аспекты популяционной генетики» (1971). Эти и многие другие работы автора в области теоретической популяционной генетики получили широкую известность. М. Кимура является членом Национальной академии наук Японии, а также ряда зарубежных академий и научных обществ.

Мы надеемся, что публикация нового большого труда М. Кимуры в русском переводе вызовет значительный интерес у наших читателей, занимающихся многообразными и увлекательными вопросами эволюции на молекулярном уровне. Книга окажется весьма полезной всем биологам широкого профиля и, несомненно, будет использоваться в качестве учебного пособия студентами, аспирантами и преподавателями.

*Ю. П. Алтухов
Л. А. Животовский*

Предисловие

В этой книге я пытаюсь убедить научный мир в том, что в основе эволюционных изменений на молекулярном уровне, т. е. изменений самого генетического материала, лежит не дарвиновский естественный отбор, а случайная фиксация нейтральных или почти нейтральных мутаций. Эта гипотеза, называемая здесь теорией нейтральности, была выдвинута мной в 1968 г. для объяснения некоторых новых открытий в области эволюции и изменчивости на молекулярном уровне и вызвала много споров. В возникшей полемике нет ничего удивительного, если учесть, что в течение более чем полувека в эволюционной биологии господствовала теория Дарвина, согласно которой живые организмы постепенно адаптируются к условиям среды путем накопления благоприятных мутаций. Естественно, эволюционисты считали этот принцип справедливым и на молекулярном уровне. Теория нейтральности не противоречит устоявшемуся взгляду на естественный отбор как на движущую силу эволюции морфо-физиологических особенностей организма, а раскрывает лишь иную сторону эволюционного процесса, подчеркивая большую роль мутаций и случайного дрейфа генов в эволюции на молекулярном уровне.

Теория нейтральности имеет два источника. Один из них — стохастическая теория популяционной генетики. Ее математический фундамент был заложен в начале 30-х годов основополагающими работами Р. Фишера, Дж. Холдейна и С. Райта. Второй источник — это молекулярная генетика, благодаря которой произошел настоящий переворот в наших представлениях о наследственной организации живого и роль которой до сих пор остается огромной. В частности, именно молекулярной генетике мы обязаны тем, что можем изучать эволюцию на уровне самой ДНК, о чем два десятилетия назад нельзя было даже мечтать.

Многие аргументы, приводимые за или против теории нейтральности, носят количественный характер: зачастую они требуют такого углубления в детали, что их приходится тщательно и подробно разъяснять. Этим, в частности, объясняется, почему объем книги получился больше, чем предполагалось вначале. Кроме того, за последние несколько лет в молекулярной биологии было сделано много новых открытий, имеющих отношение к теории нейтральности, и, конечно, я постарался по мере сил рассмотреть их.

Я с удовлетворением могу отметить, что теория нейтральности успешно выдержала испытания, выпавшие на ее долю за последние десять лет. По моему глубокому убеждению, она окажется жизнеспособной и в будущем, хотя в нее, несомненно, будут внесены дополнения и изменения. Часто говорят, что наука развивается благодаря сотрудничеству ученых всех континентов; предмет данной книги не является исключением. Читатели увидят, что в его изучении принимали участие исследователи разных стран. Даже в том случае, когда их работа предпринималась с целью опровержения теории нейтральности, она указывала пути пересмотра этой концепции, которые позволили бы сделать ее более адекватной. Теория нейтральности, как и всякая научная теория, сама эволюционировала в результате «взаимодействия» теории и эксперимента.

Создание этой книги стоило мне трех лет напряженного труда. Я испытываю особое волнение при мысли о том, что именно в 1982 г., когда она вышла в свет, отмечалась сотая годовщина со дня смерти Чарлза Дарвина. Его теория эволюции путем естественного отбора стала великим объединяющим принципом биологии. По мере появления новых данных в области молекулярной биологии перед нами открываются целые неисследованные области, которые ждут своего изучения. Надеюсь, что теория нейтральности представляет собой шаг на пути к пониманию молекулярных основ эволюции и изменчивости.

Отдельные части первоначального варианта этой книги прочли Найгел Колдер, Дэниэл Хартл, Томас Джукс, Такаси Мията, Теруми Мукаи, Масатоси Ней, Томоко Ота, Уильям Провайн, Жак Рюфье и Наоюки Такахата; их комментарии были приняты мной с благодарностью. Особенно признателен я Джеймсу Кроу за многочисленные критические замечания и помощь в преодолении семантических трудностей. Внимательно прочел всю рукопись и высказал предложения по ее улучшению К. Аюки. Пользуюсь случаем, чтобы высказать благодарность Дж. Кроу за его указания и помощь и Т. Оте за сотрудничество и конструктивную критику. Без их содействия представленная здесь теория никогда не была бы мной разработана.

Эта книга посвящается покойному профессору Таку Комаи — одному из первых эволюционных генетиков Японии. Я глубоко признателен ему за поддержку, которую он оказал мне в свое время, когда я только начинал работать в области математической генетики. Хочется думать, что, будь он жив, он встретил бы эту книгу с удовольствием и одобрением.

Введение

Теория нейтральности утверждает, что большинство эволюционных изменений на молекулярном уровне, выявляемых при сравнительном изучении аминокислотных последовательностей белков и нуклеотидных последовательностей ДНК, обусловлено не дарвиновским отбором, а случайным дрейфом селективно нейтральных или почти нейтральных мутаций. Эта теория не отрицает роли естественного отбора в определении направления адаптивной эволюции, однако она предполагает, что адаптивную природу имеет лишь незначительная часть эволюционных изменений первичной структуры ДНК, тогда как громадное большинство фенотипически «молчащих» замен нуклеотидов не оказывает никакого существенного влияния на выживание и воспроизведение и подвергается случайному дрейфу в пределах вида.

В теории нейтральности утверждается также, что большая часть внутривидовой изменчивости на молекулярном уровне, проявляющейся, например, в виде полиморфизма белков, нейтральна и поэтому большинство полиморфных аллелей, имеющих у какого-либо вида, поддерживается за счет мутационного процесса и случайной элиминации. Иными словами, теория нейтральности трактует полиморфизм белков и ДНК как переходную фазу молекулярной эволюции и отвергает представление, согласно которому большинство таких систем полиморфизма имеет адаптивное значение и в пределах вида поддерживается одной из форм балансирующего отбора.

Слово «нейтральный» используется здесь не в точном, буквальном своем значении; акцент делается не на нейтральности *per se*, а на том обстоятельстве, что основными факторами молекулярной эволюции являются мутационный процесс и случайный дрейф генов. Те мутантные гены, которые важны для осуществления молекулярной эволюции и поддержания полиморфизма, предполагаются в достаточной степени нейтральными, чтобы ведущую роль в их судьбе играл именно случай. Большое внимание в теории нейтральности уделяется селективным ограничениям, налагаемым на структуру молекул, опосредованному отбору (например, благодаря взаимосвязи меж-

ду относительным содержанием разных тРНК и частотами использования соответствующих синонимичных кодонов) и очень слабому отрицательному отбору, действующему на гены, дрейфующие к состоянию фиксации. Таким образом, теория нейтральности не содержит предположения о том, что отбор не играет никакой роли; однако в ней действительно отрицается, что сколько-нибудь заметная доля замен, происходящих в информационных молекулах, обусловлена положительным отбором и что полиморфизм этих молекул определяется силами балансирующего отбора.

Можно было бы переименовать эту теорию и назвать ее «мутационно-дрейфовой», однако термин «теория нейтральности» уже широко используется, а коней, как говорится, на переправе не меняют. Я хочу лишь объяснить читателю, что «теория нейтральности» есть сокращенное название «теории, согласно которой эволюционная изменчивость и полиморфизм на молекулярном уровне в основном обусловлены мутациями, достаточно нейтральными с точки зрения влияния на них естественного отбора, чтобы их динамика и конечная судьба определялись главным образом мутационным процессом и случайным дрейфом». Я должен также подчеркнуть, что эта теория не отрицает существования вредных мутаций. Напротив, как будет подробно показано в гл. 7, селективные ограничения, налагаемые отрицательным отбором, являются очень важной частью нейтралистской трактовки некоторых важных особенностей молекулярной эволюции.

Классическая теория эволюции с несомненностью доказала, что основным механизмом адаптивной эволюции является естественный отбор, действующий на вариации, обусловленные изменением хромосом и генов. Большую роль в нашем понимании этого процесса играют такие понятия, как численность и структура популяции, наличие экологических ниш, изменение окружающей среды, «стратегии» жизненного цикла, взаимодействие с другими видами и — в некоторых случаях — семейный отбор или, возможно, групповой отбор. В настоящее время эта область исследований значительно обогатилась благодаря выяснению неизвестных ранее молекулярных аспектов эволюции, изучение которых выявило дополнительные, совершенно новые и неожиданные, возможности и ограничения. При построении любой эволюционной теории нельзя обойтись без рассмотрения способов укладки белковых молекул с образованием третичной структуры, без выяснения того, какие из аминокислот гидрофильны, а какие гидрофобны, какие из них расположены на поверхности белковой глобулы, а какие внутри, какие связаны, а какие не связаны с жизненно важными функциями, а также без детального анализа других аспектов науки о белках, получивших мощное развитие в последнее время. Кроме того, необхо-

димо учитывать те ограничения, которые налагаются на первичную структуру ДНК особенностями вторичной структуры и спариванием оснований в молекуле РНК, наличием взаимосвязи между частотами использования синонимичных кодонов и содержанием соответствующих тРНК, процессингом РНК. Нельзя оставить без внимания совершенно новый вопрос об эволюции генетического кода, ставший актуальным после того, как было обнаружено, что словарь кодонов, используемый ДНК митохондрий, несколько отличается от обычного и к тому же варьирует. Далее, нужно учитывать новые представления об эволюции хромосом, появившиеся в результате использования новых методов их окрашивания и мечения. Необходимо помнить о том, что в будущем, по-видимому, будет обнаружено значительно больше псевдогенов — аналогов известных генов, лишенных каких-либо явных функций и выявленных с помощью методов клонирования и быстрого определения нуклеотидной последовательности ДНК. Мы должны также учитывать наличие «эгоистичной ДНК», транспозонов и других структур, ответственных за увеличение или уменьшение размера генома, а также попытаться выяснить функции так называемого «генетического хлама». Таким образом, изучение адаптивной эволюции остается все такой же увлекательной областью исследования, как и во времена Дарвина, но ведется оно на новом, значительно более высоком уровне благодаря глубокому проникновению в существо проблемы, ставшему возможным после осуществления революционных исследований в молекулярной биологии.

Однако, по моему мнению, наиболее перспективными являются молекулярно-генетические исследования, связанные с появлением концепции, согласно которой подавляющее большинство нуклеотидных замен, накапливающихся за достаточно продолжительный период времени, и преобладающая доля популяционной изменчивости полинуклеотидов в любой данный момент времени селективно нейтральны или почти нейтральны и поэтому увеличение или уменьшение частот мутантных генов обусловлено главным образом случайными процессами. Замечательно, что динамика этого процесса поддается анализу в рамках теории стохастических процессов, которая до недавнего времени считалась слишком абстрактной, чтобы иметь значение для биологии. В гл. 7 и в других главах этой книги показано, что большинство обнаруженных в последнее время явлений, перечисленных выше, подтверждают теорию нейтральности, и наиболее яркий пример такого рода представляет быстрая эволюция псевдогенов.

На случайные изменения частот нуклеотидных замен накладываются направленные, адаптивные изменения, обусловленные естественным отбором в условиях постоянного изменения среды и систематической элиминации вредных мутаций. Эти про-

цессы — объект традиционного и непреходящего интереса исследователей-эволюционистов.

Однако для создания целостной картины молекулярной эволюции необходимо учесть все типы изменений, происходящих в ДНК, каким бы слабым ни был их фенотипический эффект. Точно так же, как привлечение математического аппарата для анализа случайных процессов в физике способствовало более глубокому пониманию многих физических явлений, анализ генетических случайных процессов с помощью математических методов может пролить свет на неизвестные прежде особенности эволюции. К тому же не следует упускать из виду, что некоторые «нейтральные» аллели при соответствующих условиях среды или на другом генетическом фоне могут становиться благоприятными; таким образом, нейтральные мутации обладают скрытым потенциалом в отношении отзывчивости на отбор. Это означает, что полиморфные мутантные молекулы могут — даже если они селективно нейтральны при условиях, обычных для данного вида, — служить «сырьем» для будущей адаптивной эволюции. Что же касается трактовки случайной фиксации нейтральных мутантов как некоего «эволюционного шума», то она совершенно неадекватна и по сути дела дезориентирует исследователей. К тому же нейтральная эволюция большой совокупности генов может быть сопряжена со стабилизирующим отбором по количественным фенотипическим признакам, контролируемым большим числом сегрегирующих локусов (или сайтов) (см. разд. 6.7). В этом случае гены, замещаемые в процессе случайного дрейфа, и гены, ответственные за фенотипическую изменчивость, принадлежат к одному классу.

Здесь следует предостеречь читателя от одного возможного недоразумения. В 3-м издании словаря Уэбстера эволюция определяется как «процесс, в ходе которого любой живущий ныне организм или группа организмов приобретают через ряд изменений или эволюционных «шагов» морфологические или физиологические признаки, отличающие его [от других организмов или групп организмов]». Это определение было вполне удовлетворительным во времена Дарвина и в первой половине XX века. [Заметим, что Дарвин использовал термин «происхождение с видоизменением» («descent with modification») и высказывал мысль, что виды изменяются путем «сохранения и постепенного накопления небольших благоприятных вариаций».] Однако, если учитывать все типы изменений, выявленные методами молекулярной биологии и недоступные невооруженному глазу, то потребуются гораздо более широкое определение эволюции. В этой книге, как и в предыдущих работах по теории нейтральности, я включаю в понятие эволюции все изменения — большие и малые, видимые и невидимые, адаптивные и неадап-

тивные. В некоторых случаях эволюция может осуществляться даже путем случайной фиксации слабо вредных мутаций, коэффициенты отбора которых немного больше скоростей мутирования.

Первые две главы этой книги посвящены истории развития теорий, касающихся механизмов эволюции. Это позволяет представить теорию нейтральной молекулярной эволюции в должной исторической перспективе. Во второй части гл. 2 показано, как появилась теория нейтральности. Гл. 3 можно рассматривать как введение в эту теорию, и читателям, желающим ознакомиться сразу с существом проблемы, следует начать именно с этой главы, пропустив две первые. Однако, чтобы глубже понять излагаемую теорию, стоит все же бегло прочесть вторую половину гл. 2. В гл. 4 на примере эволюционной истории позвоночных проанализированы основные особенности эволюции на фенотипическом уровне в сравнении с особенностями молекулярной эволюции; при этом особое внимание уделяется темпам эволюции. В более общем виде вопросы молекулярной эволюции рассмотрены в следующей, 5-й, главе.

Для правильной оценки теории нейтральности необходимо иметь верное представление о сущности естественного отбора. Поэтому в гл. 6 дается его определение, рассматриваются формы и механизмы действия естественного отбора. Эта глава служит введением в классическую теорию популяционной генетики. Большая часть гл. 6 (за исключением последнего раздела) посвящена детерминистской теории. Меня удручают многочисленные работы, в которых естественный отбор рассматривается как своего рода панацея и привлекается для объяснения практически любого аспекта эволюции и изменчивости. Нетрудно придумать селекционистское объяснение почти любому частному факту; другое дело — представить соответствующие доказательства. Избежать издержек поверхностного объяснения можно только путем применения количественных подходов. В этом как раз и заключается одно из соображений, побудивших меня написать гл. 6. В гл. 7 показано, каким образом нейтралистская парадигма может привести в систему разнообразные эмпирические факты, полученные при изучении молекулярной эволюции. В этой главе подробно рассматриваются некоторые новые данные в области молекулярной биологии и генетики, представляющие дополнительные свидетельства в пользу теории нейтральности. Я удовлетворен тем, что никакая другая из существующих теорий не может дать им лучшего и более последовательного объяснения. Гл. 7 вместе с следующей главой образуют ядро этой книги. В гл. 8 довольно детально рассматриваются стохастическая теория популяционной генетики и различные модели, созданные для анализа эволюции и изменчивости на молекулярном уровне. Большинству читателей эта глава — особен-

но начиная с разд. 8.3 — может показаться трудной. Это совершенно естественно: сложен сам рассматриваемый предмет, хотя я сделал все от меня зависящее, чтобы изложить этот материал с предельной ясностью. Читатели, интересующиеся преимущественно биологическими аспектами теории нейтральности (а не математическими свойствами различных моделей), могут, прочитав первые два раздела этой главы, сразу перейти к гл. 9. В ней рассматривается вопрос о механизме поддержания генетической изменчивости на молекулярном уровне — одной из важнейших проблем, стоящих перед популяционной генетикой. Здесь я еще раз покажу, что теория нейтральности не только выдержала испытание при проверке ее с помощью различных тестов, но и оказалась полезной для объяснения уровней гетерозиготности, наблюдаемых при тех или иных условиях (например, в случае гаплоидии и диплоидии; мономерных, димерных или тетрамерных ферментов и т. п.). Проблема внутривидовой изменчивости на молекулярном уровне, можно сказать, по существу получила решение в рамках теории нейтральности.

В последней главе (гл. 10) подводится итог всей книге и сделан ряд заключительных замечаний.

Я удовлетворен тем, что теория нейтральности — в том виде, как она изложена в настоящей книге, — корректна во всех основных деталях, хотя она несомненно будет совершенствоваться в будущем (точно так же, как она постоянно совершенствовалась в прошлом). Окончательным судьей, который решит, насколько адекватна теория нейтральности, будет время, а новые данные накапливаются так быстро, что ответ на этот вопрос может быть получен очень быстро.

От Ламарка к популяционной генетике

В начале XIX в. господствовало мнение, что все живые существа на Земле суть неизменные божественные творения. Полной противоположностью этим взглядам были представления французского натуралиста Жана Батиста Ламарка. Он, напротив, считал, что живые существа могут подвергаться весьма существенным изменениям, и предложил очень простой механизм возникновения таких изменений. Ламарк опубликовал свою «Философию зоологии», в которой излагалась его концепция, в 1809 г., когда ему было уже шестьдесят пять лет. Вначале эта книга, по-видимому, не привлекла к себе особого внимания. Есть свидетельства, что в старости ослепший Ламарк был забыт своими современниками и умер в нищете. Имя его получило известность лишь много лет спустя, когда эволюция стала объектом интенсивных исследований в связи с появлением книги Дарвина «Происхождение видов». Основываясь на идеях Ламарка, несколько натуралистов из числа противников теории Дарвина создали учение, названное неоламаркизмом; роль главного фактора эволюции отводилась в нем прямому влиянию среды. (Читателям, желающим ознакомиться с авторитетным источником, который освещает подходы к изучению эволюции, использовавшиеся в XIX в., мы советуем обратиться к книге Симпсона [515].)

Наш интерес к Ламарку вызван главным образом тем, что он построил — вероятно, впервые за всю историю биологии — общую теорию эволюционного процесса. Как известно, он предположил, что результат употребления или неупотребления различных органов в соответствии с условиями жизни животного наследуется потомками; это и лежит в основе тенденции к усовершенствованию, характерной для процесса эволюции. Например, предки жирафы, пытаясь дотянуться до листьев на высоко расположенных ветвях, постоянно вытягивали ноги и шею. В результате таких «упражнений», выполняемых многими поколениями животных, соответствующие части тела постепенно удлинялись, а сама эта тенденция передавалась потомству. Так,

согласно этому воззрению, появилась современная жирафа. Таким образом, теория Ламарка основана на допущении о наследовании приобретенных признаков, сформировавшихся вследствие употребления или неупотребления органов.

Теперь ясно, что приобретенные признаки не наследуются, и поэтому теория эволюции Ламарка ошибочна. Как заметил Симпсон [515], остается только сожалеть, что такая прекрасная теория оказалась неверной.

Открыто заявил о своем несогласии с гипотезой Ламарка Август Вейсман. В серии работ он показал, что эта гипотеза маловероятна, а доказательства, используемые для ее обоснования, неубедительны. (Читателям, желающим получить представление о роли Вейсмана в истории генетики, можно порекомендовать книгу Стёртеванта [533].) Хорошо известна история о том, как Вейсман отрубал хвосты у мышей в течение двадцати двух последовательных поколений и не заметил, чтобы их длина хоть сколько-нибудь уменьшилась. Все сомнения относительно возможности наследования приобретенных признаков окончательно рассеялись в связи с интенсивным развитием менделевской генетики, завершившимся созданием современной молекулярной генетики. Тем не менее даже в наши дни еще находятся эволюционисты — приверженцы гипотезы Ламарка, утверждающие, что современная теория эволюции, основанная на менделевской генетике, неудовлетворительна. Гипотеза эта регулярно возрождается, и последнее её такое возрождение было связано с открытием обратной транскриптазы (см., например, [530]).

Вейсман решительно повел возглавляемую им школу «неодарвинизма» на бескомпромиссную борьбу с неоламаркизмом; его теоретические работы вызвали самую острую дискуссию в ранний последарвиновский период и чрезвычайно оживили эволюционные исследования в конце девятнадцатого столетия. Вейсман был также горячим поборником селекционизма, решительно защищавшим теорию отбора; говорили, что он был большим дарвинистом, чем сам Дарвин [513]. По существу, однако, Вейсман отвергал в дарвиновской теории эволюции все, кроме учения о естественном отборе. На самом деле исследование эволюции с научных позиций началось именно с работ Чарлза Дарвина. Он опубликовал свою книгу «Происхождение видов» в 1859 г., когда ему было пятьдесят лет [72], через полвека после появления «Философии зоологии» Ламарка. Мастерское владение пером и множество примеров позволили Дарвину не только убедить научный мир в том, что эволюция действительно происходила, но и показать с помощью теории естественного отбора, почему адаптивная эволюция является неизбежной. «Происхождение видов» оказало громадное влияние не только на биологию, но и на человеческую мысль в целом. Мы обяза-

ны Дарвину нашим научным пониманием природы живых существ, включая нас самих; без интеллектуальной революции, совершенной Дарвином, наша цивилизация была бы намного беднее, даже если бы в своем экономическом развитии она находилась на таком же уровне, как сейчас. В публикации, посвященной столетней годовщине выхода из печати «Происхождения видов», Г. Дж. Мёллер [391] заметил, что эту книгу можно по справедливости считать величайшей из книг, когда-либо написанных одним человеком.

Основываясь на рассмотрении искусственного отбора, который оказался очень эффективным при выведении новых сортов культурных растений и пород домашних животных, Дарвин доказывал, что такие же законы действовали в ходе эволюции в природе. Поскольку обычно рождается больше особей каждого вида, чем может выжить, между ними происходит борьба за существование, и поэтому всякое сколь угодно малое изменение, хоть в каком-то отношении выгодное данной особи, будет увеличивать вероятность ее выживания. Благодаря «закону наследственности» такое изменение имеет тенденцию к распространению. Дарвин писал:

«Как бы медленно ни совершался процесс отбора, если слабый человек мог достигнуть таких значительных результатов путем искусственного отбора, то я не вижу предела для тех изменений, той красоты и сложности взаимных приспособлений организмов друг к другу и к физическим условиям их жизни, которые могли быть осуществлены в течение долгого времени в силу естественного отбора...»

Дарвин придавал большое значение накоплению небольших полезных изменений, которое, по его мнению, лежит в основе постепенного и непрерывного процесса адаптивной эволюции.

Когда Дарвин создавал свою теорию, механизм наследственности и природа наследственной изменчивости были неизвестны, и поэтому он не мог с полной уверенностью утверждать, что естественный отбор играет ту роль, которая ему отводится в теории. В следующих один за другим переизданиях «Происхождения видов» он все с меньшей категоричностью настаивал на своем тезисе о главенствующей роли естественного отбора в эволюции, допуская, что наследование приобретенных признаков также играет важную роль. Трудно поверить, что когда-то дарвиновское учение вызвало мощную волну оппозиции и острую критику, поскольку теперь оно утвердилось настолько прочно, что считается почти неприкосновенным.

С возрождением менделевской генетики в нашем столетии открылся путь к выяснению механизма наследственности и природы наследуемых изменений, которые Дарвин столь тщетно

пытался понять. Начало менделевской эры было, однако, весьма бурным. Оно ознаменовалось острым конфликтом между «биометриками», возглавляемыми Карлом Пирсоном и У. Ф. Р. Уэлдоном, и менделистами, предводительствуемыми Уильямом Бэтсоном (см. [477]). Конфликт между двумя этими группами фактически начался еще до повторного открытия законов Менделя в 1900 г. Уэлдон, по специальности биолог, под влиянием Френсиса Гальтона пришел к убеждению, что эволюцию лучше всего изучать с помощью статистических методов. Он проделал множество измерений признаков животных и растений, с тем чтобы оценить скорость эволюции и интенсивность естественного отбора. Карл Пирсон, выдающийся специалист в области прикладной математики, заинтересовался проблемами эволюции в ходе личных контактов с Уэлдоном. Хотя теория наследственности, сформулированная Пирсоном, была ошибочна, разработанные им статистические методы, такие, например, как метод χ^2 , оказались, по мнению Холдейна [162], чрезвычайно ценными для последующего изучения эволюции и изменчивости. И Уэлдон, и Пирсон, следуя Дарвину, считали, что эволюция — постепенный процесс, протекающий под действием естественного отбора, «улавливающего» очень небольшие различия.

С другой стороны, Уильям Бэтсон, основываясь на результатах собственных исследований изменчивости растений и животных, пришел к вполне определенному выводу, что, вопреки мнению Дарвина, эволюция не могла происходить с помощью естественного отбора, действующего на непрерывную изменчивость. В этой связи Бэтсон указывал на важное значение дискретной изменчивости.

После «переоткрытия» законов Менделя конфликт между менделистами и биометриками еще более обострился. В то время как Бэтсон находился под большим впечатлением от этого события, понимая всю важность законов Менделя, Уэлдон и Пирсон занялись энергичной критикой менделизма. Рассказывают, что Уэлдон бросил все свои силы на опровержение менделизма и в поисках исключений из этих законов пересмотрел множество огромных томов заводских племенных книг породистых лошадей, подорвал свое здоровье, заболел пневмонией и умер молодым (см. [464]). Как один из лауреатов, удостоенных Уэлдоновской мемориальной премии, я испытываю особый интерес к жизни Уэлдона и хотел бы добавить, что он сделал важное открытие, касающееся действия естественного отбора. Изучая улиток [578], он подсчитывал у них число витков, приходящихся на заданный отрезок оси раковины. Сравнивая это число для молодых и взрослых особей, он обнаружил, что молодые улитки, у которых число витков слишком мало или слишком велико, характеризуются более высокой смертностью, чем улитки, раковины которых близки к средним. Это было, как

отметил Холдейн [165], одно из первых сообщений относительно «центростремительного отбора».

После смерти Уэлдона биометрики постоянно отступали, и вскоре победа менделистов, опиравшихся на многочисленные факты, стала очевидной. В это время многие менделисты сомневались, что естественный отбор, оперирующий небольшими непрерывными изменениями, мог, в согласии со взглядами Дарвина, быть эффективным фактором эволюции. Более того, они взяли на вооружение мутационную теорию Гуго де Фриза, утверждая, что новые виды возникают посредством мутационных скачков, а не в результате постепенного действия естественного отбора. В то время мутационная теория, предложенная в начале нашего века, стала очень популярной у биологов и приобрела множество сторонников.

Теперь известно, что «мутации», наблюдавшиеся де Фризом у ослинника *Oenothera lamarckiana*, вероятно, были обусловлены гетерозиготностью по особым множественным перестройкам хромосом. Как заметил Стёртевант [533], есть какая-то ирония в том, что теперь лишь немногие из этих «мутаций» можно было бы назвать мутациями. Тем не менее теория де Фриза, привлекая широкое внимание к мутациям как возможной причине генетической изменчивости, способствовала уточнению понятия мутации, окончательно утвердившемуся благодаря Г. Дж. Мёллеру, громадный вклад которого в изучение механизма эволюции будет освещен ниже.

В первое десятилетие XX в. были поставлены целенаправленные эксперименты для выяснения вопроса о том, эффективен естественный отбор по количественным («непрерывным») признакам (как полагал Дарвин) или нет. Вероятно, самые известные из этих экспериментов провел Вильгельм Иоганнсен. Он предложил теорию чистых линий, показав, что отбор в чистой линии неэффективен. Богатая событиями и противоречивая атмосфера этих лет отражена в прекрасном историческом очерке Провайна [477].

Постепенно, однако, стало ясно, что менделизм и дарвинизм вполне совместимы друг с другом. Такая перемена взглядов была обусловлена интенсивным развитием генетики, причем особое значение приобрела генетика дрозофилы, изучение которой показало, что мутационные изменения могут быть очень небольшими. В конечном счете попытки осуществить синтез дарвинизма и менделизма с помощью методов биометрии привели к созданию популяционной генетики.

Удобным исходным пунктом при изложении истории популяционной генетики является результат, полученный Харди (1908 г.) [168] и Вайнбергом (1908 г.) [576]. Эти авторы показали, что при случайном скрещивании и менделевском наследовании частоты генотипов по данному аутосомному локусу из по-

коления в поколение остаются неизменными. Кроме того, они показали, что если два аллеля, A и a , встречаются с относительными частотами p и q ($p+q=1$), то равновесные частоты генотипов будут равны

$$p^2AA : 2pqAa : q^2aa.$$

Вывод о постоянстве генотипических частот имел важное значение, так как в то время некоторые биометрики критиковали законы Менделя, утверждая, что если бы они были справедливы, то частота доминантного признака (такого, например, как брахидактилия у человека) увеличивалась бы до 75%. Работа Харди помогла рассеять это заблуждение. В настоящее время, когда законы Менделя твердо установлены, у нас нет необходимости заниматься опровержением подобных умозаключений.

Одним из аспектов обобщения Харди и Вайнберга, до сих пор не потерявшим актуальности, является правило, согласно которому частоты генотипов (или зигот) в случайно скрещивающейся популяции можно получить простым перемножением соответствующих частот генов. Тем не менее сам закон Харди—Вайнберга — не более чем полезное правило. Обобщение, сделанное этими авторами, служит лишь удобным исходным пунктом при изложении основ популяционной генетики, и мне кажется странным, что его значение обычно сильно преувеличивают. Такая традиция, вероятно, идет от Добржанского, который, излагая «закон Харди—Вайнберга» в третьем издании (1951 г.) своей авторитетной книги «Генетика и происхождение видов», писал о нем как об основополагающем принципе популяционной генетики и современной эволюционной теории. По моему мнению, принятый ныне педантичный, не способный пробудить живую мысль способ изложения закона Харди—Вайнберга не может принести ничего, кроме вреда, и поэтому преподавание элементарной популяционной генетики нуждается в серьезной реформе. Во многих учебниках закон Харди—Вайнберга после его торжественного провозглашения важнейшим законом популяционной генетики формулируется следующим образом: 1) в отсутствие возмущающих факторов (таких, как мутации, отбор, миграция, случайная выборка гамет и т. д.) и при случайном скрещивании частоты генотипов остаются неизменными; 2) эти частоты равны p^2AA , $2pqAa$ и q^2aa . Кроме того, подчеркивается, что в первом пункте сформулирована более важная часть закона, чем во втором, а это, на мой взгляд, является чистым анахронизмом.

То, что частоты генов не изменяются в отсутствие возмущающих факторов, само собой разумеется, поскольку гомологичные гены регулярно расщепляются в мейозе и все гены являются самовоспроизводящимися единицами. Мы нередко забываем, что образование генотипов (зигот) есть просто объедине-

ние двух гомологичных генов. Так уж сложилось, что у высших организмов, в том числе и у человека, главной фазой жизненного цикла является диплоидная фаза; вероятно, никто не обратил бы внимания на закон Харди — Вайнберга, если бы мы, люди, были гаплоидными организмами. Закон Харди — Вайнберга не представлял бы никакого интереса, если бы он сводился только к правилу 1. С другой стороны, в качестве приближенного метода расчета этот закон действительно полезен; если мы произведем подсчет генов и генотипов сразу после оплодотворения, то сможем с хорошей точностью оценить генотипические частоты, просто перемножив соответствующие генные частоты, даже если они изменяются из поколения в поколение под давлением естественного отбора.

В течение двух десятилетий, прошедших после опубликования работ Харди и Вайнберга, Р. Фишер, Дж. Холдейн и Сьюэлл Райт получили основные популяционно-генетические следствия из законов Менделя. В начале 30-х годов математическая теория классической популяционной генетики в сущности была завершена.

Из этих трех основателей теоретической популяционной генетики, по-видимому, именно Фишер оказал наибольшее влияние на формирование ортодоксального взгляда, согласно которому темпы и направление эволюции определяются почти исключительно естественным отбором, а мутации, миграция и случайный дрейф играют второстепенную роль.

Ортодоксальную эволюционную генетику часто называют неodarвинизмом, так как она развивается в русле традиции, идущей от Вейсмана, который отрицал наследование приобретенных признаков и придавал первостепенное значение естественному отбору. Широко используется также более выразительное название «синтетическая теория эволюции», вероятно, с целью подчеркнуть, что эта концепция основывается на рассмотрении различных факторов эволюции.

В области теоретической популяционной генетики много крупных результатов получил Фишер. В 1918 г. он опубликовал фундаментальную работу, посвященную анализу корреляций между родственниками с помощью статистических методов, основанных на предположении о менделевском наследовании [104]. Как и во многих других его трудах, изучаемый вопрос был разработан столь детально и получил такое убедительное математическое обоснование, что все дальнейшие исследования, выполненные целым рядом специалистов по математической генетике, смогли внести лишь относительно небольшой вклад в разработку данного вопроса. Даже спустя более чем полвека эта работа фактически все еще оставалась самым фундаментальным исследованием в области биометрической генетики. В 1922 г. Фишер опубликовал статью с довольно странным на-

званием «О доминантном отношении» [105]. Хотя математическая трактовка проблемы была весьма несовершенной и даже содержала серьезные ошибки, эта работа чрезвычайно оригинальна: она послужила началом использованию в популяционной генетике стохастических методов, с помощью которых можно изучать случайные флуктуации генных частот, происходящие при смене поколений. В качестве фактора, вызывающего случайную флуктуацию частоты гена, Фишер рассматривал процесс случайной выборки гамет; исключив отбор, он изучал влияние этого фактора (названное им «эффектом Хагедорна») на уменьшение генетической изменчивости вида. Используя частоту гена, преобразованную с целью стабилизации дисперсии, он получил дифференциальное уравнение в частных производных такого же вида, как уравнение теплопроводности. Фишер решил его и установил, что дисперсия уменьшается со скоростью $1/(4N)$ за поколение, где N — число скрещивающихся особей. Отсюда Фишер сделал вывод, что уменьшение изменчивости, обусловленное этим фактором, — процесс чрезвычайно медленный, а его влияние на эволюцию незначительно и, напротив, даже самый слабый отбор имеет существенное значение. К сожалению, работа Фишера содержала несколько ошибок. В частности, скорость уменьшения дисперсии на самом деле, как показал Райт, использовавший совершенно другой метод, равна $1/(2N)$. Позднее Фишер [106] проанализировал эту проблему более детально, исправив допущенные ошибки и существенно дополнив результаты, полученные ранее. Однако он не нашел оснований для изменения своего прежнего заключения о роли процесса случайной выборки гамет; по его мнению, численность популяций большинства видов столь велика, что процесс случайной выборки — фактор, для эволюции совершенно несущественный. В работе Фишера, опубликованной в 1922 г. [105], сообщается также об одном открытии, оказавшем глубокое влияние на последующее развитие научной мысли в области популяционной и эволюционной генетики: Фишер показал, что если отбор благоприятствует гетерозиготам, то существует некоторое состояние устойчивого равновесия и в популяции сохраняются оба аллеля. Иными словами, сверхдоминантные аллели активно поддерживаются естественным отбором.

Работа Фишера, опубликованная в 1930 г. [106], посвящена дальнейшей разработке «стохастической проблемы», поставленной в статье [105]. В ней он предложил остроумный подход с использованием метода функциональных уравнений, позволяющий прогнозировать судьбу редких мутантных аллелей в популяции. Хотя теперь этот метод хорошо известен математикам, изучающим ветвящиеся процессы, для того времени уровень сложности был чрезвычайно высоким. Фишер так глубоко проник в сущность изучаемой проблемы и с таким совершенством

владел математическим аппаратом, что, по моему мнению, превзойти его в математическом плане будет нелегко. Достигнутые нами успехи связаны главным образом с непредвиденной устойчивостью и универсальностью метода диффузионных уравнений (называемых теперь «диффузионными моделями»; см. [228]); не располагая теми знаниями, какие имеются теперь в нашем распоряжении, Фишер не мог не испытывать определенных сомнений, применяя этот метод для изучения поведения редких мутантных аллелей. Так или иначе, он пришел к заключению, что случайные процессы, а точнее — случайный дрейф частот генов, в эволюции роли не играет. Лично мне после того, как я ознакомился с его впечатляющим математическим обоснованием проблемы стохастического поведения мутантов, этот вывод показался нелогичным.

Целостная концепция эволюционного процесса, разработанная Фишером, была изложена им в книге «Генетическая теория естественного отбора» [107], которая внесла большой вклад в осуществление синтеза дарвинизма, менделизма и биометрии, чем любая другая.

Эта книга стала своего рода библией для многих биологов, пытавшихся понять эволюцию с позиций теории естественного отбора. Как и библия, она явилась неким откровением и в то же время содержала ряд тенденциозных идей. Поэтому именно Фишер в значительной мере способствовал последующему широкому распространению панселекционизма в Англии.

Одна из его любимых теорий (теория доминантности) дает объяснение процессу эволюции доминантности. Фишер выдвинул предположение, что доминирование гена дикого типа над мутантными аллелями обусловлено продолжительным отбором модификаторов, которые придают гетерозиготам по мутантным аллелям сходство с гомозиготами дикого типа [107]. Он полагал, что вначале приспособленность гетерозигот по мутантным аллелям является промежуточной между приспособленностями двух гомозигот (т. е. никакого доминирования нет). Так как гетерозиготы по мутантным аллелям обычно редки, отбор соответствующих модификаторов чрезвычайно слаб (характеризуется величинами того же порядка, что и частота мутирования); тем не менее Фишер считал, что при длительной эволюции этот очень слабый отбор в конечном счете окажется эффективным. Для количественной оценки эффективности отбора он использовал выведенную им формулу, которая выражает скорость прогресса, достигнутого в результате модификации гетерозиготы, по сравнению со скоростью прогресса, который был бы вызван отбором такой же интенсивности, действующим на популяцию, состоящую из одних гетерозигот. Эта формула с использованием обозначений и терминов, принятых в данной книге, имеет вид

$$2\sigma w_1 s / (1 - w_1 s) (2 - w_1 s), \quad (1.1)$$

где v — скорость мутирования, ω_{12} — относительная приспособленность гетерозигот. Против фишеровской теории доминантности выступал Райт (например, [588]; подробное изложение этого вопроса см. в гл. 15 3-го тома фундаментальной сводки Райта [603]). Он указал, что столь слабый отбор был бы легко сведен на нет другими силами отбора, которые могли быть вызваны влиянием модификаторов на гораздо более многочисленные гомозиготы дикого типа, и к тому же был бы «снят» мутационным давлением и случайным дрейфом. Критиковал теорию Фишера и Холдейн [155]; он считал, что гораздо более вероятно такая ситуация, когда отбор действует на аллели дикого типа, благоприятствуя более активным и, следовательно, более доминантным их формам.

В настоящее время представляется, что теория доминантности Фишера либо неверна совсем, либо в лучшем случае применима только к особым ситуациям. Недавние исследования летальных и слабо вредных мутаций у дрозофилы ясно показали [380], что в среднем для слабо вредных мутаций характерна гораздо более высокая степень доминантности (0,3—0,4), чем для летальных и полублетальных (0,01—0,02). В цитируемой работе изучалась только выживаемость, однако тот же самый вывод справедлив и в отношении общей приспособленности. Мутации с умеренным эффектом в гомозиготном состоянии обладают значительно большим относительным влиянием на приспособленность гетерозигот, чем «радикальные» мутации [65, 510]. Из (1.1) очевидно (и это подчеркивал сам Фишер), что, чем ближе приспособленность гетерозиготы к приспособленности гомозиготы по нормальным аллелям, тем быстрее будет расти скорость модификации доминантности. Поэтому следует ожидать, что умеренно вредные мутации будут ближе к почти полностью рецессивному состоянию, чем летали. Однако на самом деле это не так. С подобной критикой теории Фишера выступал Чарлзуорс [45] (см. также [46]). Более правдоподобными, чем теория Фишера, представляются гипотезы Райта [603] и Мёллера [386]. По мнению этих исследователей, аллель дикого типа обычно детерминирует синтез какого-то полезного соединения, которое рецессивный мутант или не образует совсем, или образует в меньших количествах. Таким образом, феномен доминантности можно понять на основе анализа зависимости «доза — эффект». В процессе эволюции гены дикого типа подвергаются отбору на надежность в отношении своей способности продуцировать полезные вещества, что позволяет им противостоять действию неблагоприятных факторов — внешних и внутренних; доминантность же оказывается автоматическим следствием такого отбора.

Во всех своих трудах Фишер все время подчеркивал, что роль, которую играл в эволюции случайный дрейф, минималь-

на, и это не могло не повлиять на взгляды молодых английских генетиков, оставивших без внимания эту проблему. Оказалось, однако, что основные успехи теоретической популяционной генетики, достигнутые за последние годы, связаны именно с анализом изменений частот генов как стохастических процессов [308]. Отсутствие интереса к этой области исследования со стороны английских генетиков, по всей вероятности, обусловлено влиянием Фишера. Совершенно иная ситуация сложилась в Шотландии, где Алан Робертсон из Института генетики животных, испытавший влияние Райта, выполнил выдающееся исследование, посвященное изучению малых популяций.

Работы Холдейна в области популяционной генетики менее оригинальны, чем работы Фишера, однако они более разнообразны, в известном смысле более объективны и во многих случаях более приемлемы в биологическом плане. В 1924 г. он приступил к публикации серии статей под общим названием «Математическая теория естественного и искусственного отбора». Эти статьи положили начало систематическому математическому изучению динамики генных частот под действием естественного отбора. В основе его работы лежит убеждение, что удовлетворительная теория естественного отбора должна быть количественной (это положение было сформулировано в первой статье указанной серии [153]) и что адекватность генетической теории естественного отбора можно проверить только с помощью количественных исследований. Холдейн рассматривал такие разновидности отбора, как зиготический, гаметический, семейный, а также отбор по X-хромосоме. Типичную ситуацию, изучавшуюся им, можно описать так. В очень большой популяции со случайным скрещиванием отбор благоприятствует доминантным особям (AA, Aa), а из рецессивных особей (aa) до репродуктивного возраста доживает, по сравнению с доминантными, $(1-k)$ -я часть; величину k называют коэффициентом отбора, благоприятствующего аллелю A . Холдейн вывел уравнение в конечных разностях, с помощью которого можно перейти от генного отношения (т. е. отношения частот A и a) для данного поколения к отношению для следующего поколения. Это уравнение нелинейно, и хотя Холдейн не смог получить его точное решение, он нашел приближенные решения, вполне приемлемые для практических целей. Далее он составил таблицу, позволяющую легко установить связь между интенсивностью отбора и временем изменения генотипических частот. Например, если $k=0,001$, то для увеличения частоты A от 1 до 99% потребуется около 16 483 поколений. Если k будет в десять раз больше, то число необходимых для этого поколений составит одну десятую от приведенной величины. Холдейн показал, что число поколений, необходимых для заданного изменения частоты гена, в общем случае обратно пропорционально интенсивности отбора. Он применил эту тео-

рию к случаю индустриального меланизма у бабочек, обитающих в районе Манчестера, и установил, что коэффициент действующего здесь отбора не меньше 0,332, а возможно, достигает 0,5. Примерно тридцать лет спустя Кеттлуэлл [217] и другие исследователи установили, что обычно на этих бабочек действует отбор именно такой интенсивности и осуществляется он главным образом птицами, которые ими питаются.

Подобный анализ изменения генных частот, начало которому положил Холдейн, может быть назван детерминистическим, так как в нем не рассматриваются элементы случайности, в частности процесс случайной выборки гамет. Даже несмотря на то, что с тех пор были разработаны более тонкие методы, детерминистический анализ все еще широко применяется и обычно, когда численность популяции достаточно велика, оказывается вполне удовлетворительным. Более того, благодаря своей простоте этот подход до сих пор остается наиболее полезным и часто единственным приемлемым подходом к решению многих проблем. В последующих статьях указанной серии Холдейн рассматривал влияние на динамику генных частот различных факторов: неполного инбридинга, ассортативного скрещивания, неполного доминирования (в том числе сверхдоминирования) в аутосомных, а также в сцепленных с полом локусах, многофакторной наследственности, сцепления (в отсутствие отбора), полиплоидии, перекрытия поколений и изоляции. В этих статьях представлены результаты, до сих пор не утратившие своего значения, хотя нередко их недооценивают. Из работ данной серии особого упоминания заслуживают две: сообщения V и VIII. В сообщении V [154] Холдейн формулирует «стохастическую проблему» и изучает вероятность фиксации мутантных генов, используя метод производящих функций, предложенный Фишером [105]. Он первым показал, что доминантный мутантный ген, имеющий небольшое селективное преимущество k в большой панмиктической популяции, в конечном счете зафиксируется с вероятностью примерно $2k$. В другой его выдающейся работе — сообщении VIII из той же серии, вышедшем под названием «Метастабильные популяции» [156], — рассмотрен случай, когда мутантные гены сами по себе являются вредными, но в сочетании друг с другом приобретают селективные преимущества. Предположив для простоты, что сцепление отсутствует, Холдейн с редким изяществом проанализировал ситуацию с двумя локусами, определив траектории точек, представляющих генетический состав популяции, в двумерном пространстве. Рассматривая превращение популяции $AABB$ в популяцию $aabb$, он писал, что вид, претерпевающий превращение такого типа, можно называть «метастабильным». Он предположил, что процесс видообразования сводится к нарушению метастабильного равновесия и что такое отклонение от равнове-

сия особенно вероятно для небольших изолированных сообществ. Весьма примечательно сходство этого вывода с теорией Райта, предложенной независимо от Холдейна и трактующей эволюцию как процесс проб и ошибок, который рассматривается в рамках концепции многомерной адаптивной поверхности («теория смещающегося равновесия», принадлежащая Райту, будет обсуждаться позднее).

Результаты Холдейна, полученные им к этому времени, подытожены в его книге «Факторы эволюции», опубликованной в 1932 г. [157]. Эта книга, значительно более легкая для чтения, чем книга Фишера, вышедшая в 1930 г., внесла большой вклад в распространение идеи о взаимной совместимости менделизма и дарвинизма среди широкого круга биологов. Холдейн неоднократно высказывал свою точку зрения относительно использования математики в биологии и особенно в генетической теории эволюции. В одной из своих статей [159] он писал:

«В настоящее время существует мнение, что математическая теория эволюции оказалась в невыгодном положении, так как она слишком «математична», чтобы заинтересовать большинство биологов, и недостаточно «математична», чтобы вызвать интерес у математиков. Тем не менее есть все основания полагать, что в следующей половине нашего столетия она разовьется в самостоятельный раздел прикладной математики».

Его предсказание полностью сбылось, и теперь математические модели в биологии находят все более широкое применение. Убежденность Холдейна в огромном значении перехода эволюционных исследований на более высокий количественный уровень нашла яркое отражение в его блестящей статье под названием «Слово в защиту генетики мешка с бобами» [166]. Он написал ее незадолго до своей смерти, отвечая на критику Эрнста Майра [353], который утверждал, что в теории популяционной генетики популяция рассматривается как мешок, наполненный разноцветными бобами, и что такое упрощенное теоретизирование приводит к ложным представлениям.

В отличие от Фишера и Холдейна Райт в своей концепции эволюции придает большое значение случайному генетическому дрейфу и системе неаддитивного взаимодействия генов. Развивая идеи, заложенные в своей ранней работе о системах скрещивания [585], в которой он предложил и разработал новый, очень эффективный подход к изучению инбридинга и ассортативного скрещивания (так называемый метод коэффициентов путей), Райт опубликовал в 1931 г. статью под названием «Эволюция в менделевских популяциях» [586]. Эта статья наряду с трудами Фишера «Генетическая теория естественного отбора» [107] и Холдейна «Факторы эволюции» [157] явилась кульми-

национным моментом в развитии классической популяционной генетики, осуществив полный синтез дарвинизма и менделизма.

Начиная с 1932 г. Райт развивал теорию эволюции, названную им позднее «теорией смещающегося равновесия» [587, 602]. Изучая различные «модусы трансформации» под действием мутационного давления (что может наблюдаться при редукции бесполезных органов), массового отбора, миграции и дрейфа, обусловленного только случайностями выборки, Райт пришел к выводу, что наиболее подходящими для осуществления быстрого эволюционного прогресса в результате смещения равновесия являются большие подразделенные популяции. Согласно Райту [603], этот процесс состоит из трех стадий: 1) случайного дрейфа генов, когда в каждом деде (локальной популяции) происходят значительные флуктуации генных частот, обусловленные случайностями выборки или флуктуациями параметров, связанных с давлением различных эволюционных факторов; 2) массового отбора, когда тот или иной дем случайно пересекает одну из бесчисленных «двухфакторных седловин» на поверхности приспособленностей, что приводит к быстрому генетическому изменению этого деда при массовом отборе; 3) междемового отбора, когда дем, достигший нового пика приспособленности, более высокого по сравнению с теми, на которых находятся соседние деды, увеличивается в результате междемового отбора. Этот процесс может распространяться концентрическими кругами, идущими из разных центров, и два таких круга, перекрываясь, могут привести к образованию нового, еще более высокого пика, который послужит исходным пунктом дальнейшего прогресса. Таким образом, поистине бесконечное пространство взаимодействующих генетических систем может быть «обследовано» при небольшом числе новых мутаций.

Теорию Райта часто трактуют неправильно, полагая, что случайный дрейф рассматривается в ней как важнейшая альтернатива естественному отбору в фенотипической эволюции видов. Такое представление ошибочно и не соответствует содержанию теории, на что указывал сам Райт (см. с. 453 в т. 3 фундаментального труда Райта [603]).

Согласно Райту, в основе эволюции лежит процесс смещения равновесия, в который вовлечены все эволюционные факторы, действующие при этом совместно. В течение одного-двух десятилетий после опубликования его статьи, вышедшей в 1931 г. [586], эволюционная теория Райта привлекла к себе всеобщее внимание. Популярности этой теории в значительной мере способствовало подробное изложение ее в книге Добржанского «Генетика и происхождение видов» [82], которая была очень известна среди биологов. Я до сих пор помню то гипнотическое состояние, в котором я находился, когда, будучи еще студентом, впервые прочитал о теории Райта в книге Добржанского.

Райта критиковали Фишер и представители его школы, сложившейся в Англии [109, 110, 120, 598]; эта дискуссия продолжается и поныне. «Играет ли случайный дрейф какую-либо важную роль в эволюции?» — так формулируется вопрос, вокруг которого ведется дискуссия. Фишер был уверен, что число особей, составляющих данный вид, обычно так велико, что влияние случайных событий, связанных с процессом выборки гамет, пренебрежимо мало. Он также считал маловероятным, что значения произведений численности популяции на коэффициенты отбора для большинства мутантных аллелей остаются в ходе эволюционного процесса в окрестности нуля, т. е. полагал, что селективно нейтральные мутанты должны быть крайне редкими. С другой стороны, как уже отмечалось, Райт признавал важное значение случайного дрейфа в локальных популяциях для прогрессивной эволюции. Он считал, что вид, представляющий собой одну большую панмиктическую популяцию, посредством массового отбора быстро достигнет адаптивного плато и придет к эволюционному застою. Фишер же рассматривал такое событие как маловероятное [109], поскольку популяция содержит громадное число генотипов, и если какой-нибудь крайний из них (гомозиготный) обладает в какой-либо локальности селективным преимуществом, «то ничто не сможет предотвратить сдвига генных отношений к значениям, обеспечивающим более высокий уровень адаптации».

Теория Райта встретила возражение и у Майра ([353]; см. с. 520), который считает, что у широко распространенных преуспевающих видов существует значительная миграция между демами, приводящая к сближению частот генов во всех локальных популяциях. По моему мнению, никаких четких доказательств правильности теории смещающегося равновесия не получено, хотя эта гипотеза остается весьма привлекательной. Кроме того, мне представляется довольно удивительным, что, несмотря на свое страстное увлечение проблемами, связанными с процессом смещающегося равновесия, Райт не провел широкого количественного анализа этого процесса и не предпринял серьезных попыток проверить свою теорию, с тем чтобы опровергнуть аргументы критиков.

Хотя теория эволюции Райта весьма противоречива, его работы, опубликованные после 1931 г., внесли значительный вклад в наше понимание стохастического поведения мутантных аллелей в популяциях конечной численности. Он получил распределения генных частот в популяциях при наличии как прямых, так и обратных мутаций ([590, 591, 593, 595], см. обзор в т. 2 фундаментального труда Райта [601]). Последующая разработка этой тематики с помощью диффузионных моделей (см. [228]) в значительной мере основывается на результатах, полученных Райтом.

Развитие синтетической теории эволюции и рождение теории нейтральности

2.1. Становление синтетической теории и ее превращение в ортодоксальную концепцию эволюционного процесса

В предыдущей главе был дан очерк истории развития теоретических представлений о механизме эволюции; этот процесс завершился в начале 30-х годов созданием Фишером, Холдейном и Райтом классической популяционной генетики. Говоря о синтезе дарвинизма и менделевской генетики, нельзя не упомянуть еще одного имени. Речь идет о Г. Дж. Мёллере, который в начале 20-х годов пролил свет на природу генных мутаций и их роль в эволюции путем естественного отбора.

Хорошо известно его эффектное доказательство мутагенного действия рентгеновских лучей [384]; за эту работу Мёллеру была присуждена Нобелевская премия. О его фундаментальном вкладе в наше понимание механизма эволюции известно менее широко. Мёллер предложил концепцию (убедительно обосновав ее), согласно которой в основе адаптивной эволюции путем естественного отбора лежит одно замечательное свойство гена, состоящее в том, что самовоспроизводящимся является не только сам ген, но и его мутантные формы ([383]; см. также с. 175 монографии Мёллера [392]). Поэтому естественный отбор можно представить как дифференциальное размножение мутантных форм. Мёллер пошел еще дальше, выдвинув тезис ([385], см. с. 188 его монографии [392]), что ген есть основа самой жизни. И эта мысль была высказана в то время, когда многие биологи все еще считали ген неким гипотетическим фактором, введенным для объяснения результатов скрещиваний. Позже Мёллер обобщил концепцию «примата гена» [393]. Согласно Мёллеру, *«критерием жизни является потенциальная способность к эволюции путем дарвиновского естественного отбора»*. Он также обосновал положение, согласно которому генные мутации, возникая по воле слепого случая, обычно вредны и, чем более они «радикальны», тем, как правило, более вредны. Теперь мы редко отдаем должное проницательности Мёллера, так как это представление принимается нами как само собой разумеющееся. Другая закономерность, на которую он обратил внимание, рассматривая ее как следствие из приведенного здесь положения,

состоит в том, что мутантные гены, как правило, менее доминантны, чем их «нормальные» предшественники. Все эти положения привлекались им не только для объяснения особенностей эволюционного процесса, но и для анализа проблем, касающихся физического здоровья общества. Его статья под названием «Наш груз мутаций» [388] принадлежит к числу классических работ по популяционной генетике человека. Одно из замечательных предсказаний Мёллера — предположение о том, что большинство мутантных генов характеризуется определенной степенью доминантности и что эта степень выше в случае менее вредных мутантных генов, чем в случае более вредных. Впоследствии это предположение получило полное подтверждение в работах Кроу, Мукаи и их сотрудников [65, 510].

Мёллер не только ввел эти новые тогда представления, совершив своего рода «концептуальную революцию», но и создал первоклассные специальные линии дрозофил, содержащие знаменитую хромосому «*CIB*», что сделало дрозофилу ценнейшим объектом для экспериментальных генетических исследований; в этом смысле она уступает лишь *E. coli*, которая, однако, стала использоваться гораздо позже. Без крупных методических разработок, осуществленных Мёллером, изучение летальных и вредных мутантных генов в природных популяциях дрозофил было бы, конечно, невозможно. Мёллер более чем кто-либо другой способствовал созданию фундамента, на котором зиждется изучение эволюции методами генетики.

На основе, заложенной Фишером, Холдейном, Райтом, а также Мёллером, были выполнены многочисленные исследования. Среди них особого внимания заслуживают работы Ф. Добржанского и его школы по генетике природных популяций, палеонтологические работы Дж. Симпсона, исследования Е. Форда и его учеников по «экологической генетике» и теоретические построения Эрнста Майра, касающиеся видообразования. Результаты всех этих работ привели к созданию единой синтетической теории эволюции, постепенно сформировавшейся в стройную и впечатляющую систему взглядов.

Основной вклад в популяционную генетику, сделанный Добржанским, состоит помимо тех выводов, которые содержатся в его авторитетных печатных трудах, в том, что он совместно с А. Стёртевантом обнаружил у некоторых видов дрозофил высокий уровень полиморфизма хромосом, включая инверсии (т. е. обратный порядок расположения генов в тех или иных участках хромосом) (см. обзор в 3-м издании книги Добржанского [83]). Он провел обширное и весьма успешное исследование инверсионного полиморфизма (преимущественно в природных популяциях *D. pseudoobscura*), а помимо этого предпринял лабораторные эксперименты с целью изучения динамики частот хромосом с различными инверсиями. Эта работа проводилась сов-

местно с Райтом, который оценивал приспособленности различных генотипов [605]. Статистический анализ показал, что гетерозиготы по инверсии имеют более высокую приспособленность, чем гомозиготы. Результаты экспериментов были проанализированы и в свете гипотезы частотно-зависимого отбора, также обнаружившей соответствие с этими данными. Впоследствии Добржанский и его сотрудники попытались применить более прямой тест на гетерозис по полиморфным инверсиям. Хотя наборы генов, по которым различались эти инверсии, известны не были, столь широкое распространение инверсионного полиморфизма и экспериментальные данные, свидетельствующие о выраженном гетерозисе по инверсиям, привели Добржанского к мысли о широком распространении сверхдоминирования (преимущества гетерозигот) по отдельным генным локусам.

Энтузиазм, вызванный идеей всеобщности превосходства гетерозигот, был подкреплен концепцией «генетического гомеостаза», сформулированной Лернером [299]. Согласно Лернеру, менделевские популяции (т. е. воспроизводящиеся сообщества) способны поддерживать равновесие и это приводит к сохранению их генетического состава, обеспечивающего максимальную среднюю приспособленность. Он считал, что наиболее вероятный механизм этого процесса соответствует модели с преимуществом гетерозигот и что для нормального развития особи необходим определенный минимальный уровень гетерозиготности. Он даже пришел к мысли ([299], с. 75), что доменделевское представление о вредности «инбридинга *per se*» в какой-то степени верно. Иными словами, «можно считать, что инбредная дегенерация обусловлена не только генным составом, но и гомозиготностью как таковой». Лернер возражал против концепции генетического груза, предложенной Мёллером [388], выдвигая довод, что «равновесие в популяциях, основанное на селективном преимуществе гетерозигот», — феномен, имеющий широкую сферу проявления, которая не ограничивается отдельными случаями полиморфизма. Находясь под влиянием концепции Райта, в которой эволюция рассматривается как нерегулярное смещение состояния равновесия, Лернер подчеркивал, что приспособленность зависит не только от сверхдоминирования, но и от эпистатического взаимодействия генов.

Эти представления, в которых первостепенное значение придается гетерозиготности как «адаптивной норме» и эпистазу как причине, приводящей к формированию «коадаптированных генных комплексов», получили дальнейшее развитие в выдающейся работе Добржанского [84]. Эта работа была представлена на 20-й симпозиум в Колд-Спринг-Харборе, посвященный вопросам популяционной генетики. Добржанский выдвинул две альтернативные рабочие гипотезы, одна из которых была названа «классической», а другая — «балансовой». В основе клас-

сической гипотезы лежит предположение, что эволюционные изменения сводятся к последовательным заменам и фиксации более благоприятных аллелей вместо менее благоприятных. Согласно этой гипотезе, преобладающая часть особей в популяции гомозиготна по большинству генов и, следовательно, гетерозиготные локусы составляют меньшинство. В классической гипотезе рассматриваются четыре главных источника редких гетерозигот: 1) вредные мутации, которые через определенное число генераций элиминируются естественным отбором; 2) адаптивно нейтральные мутации (*adaptively neutral mutations*); 3) адаптивный полиморфизм, поддерживаемый за счет гетерогенности среды; 4) редкие полезные мутации, постепенно распространяющиеся в популяции.

В основе балансовой гипотезы лежит предположение, что адаптивная норма представлена множеством генотипов, гетерозиготных по большому числу аллелей. В нормальных аутбредных популяциях гомозиготы по этим аллелям составляют лишь небольшую часть особей и по приспособленности уступают гетерозиготам. В данном случае должно существовать давление отбора, благоприятствующего появлению во многих локусах серий множественных аллелей. Балансовая гипотеза в изложении Добржанского очень напоминает концепцию генетического гомеостаза Лернера. Из работ Добржанского совершенно очевидно, что он считал себя сторонником балансовой гипотезы, а Мёллера отнес к противостоящему лагерю «классиков».

Для подкрепления своей точки зрения Добржанский представил результаты различных экспериментов [84]; самый замечательный из них заключался в получении посредством рекомбинации так называемых синтетических леталей. В одном из опытов, о которых он докладывал, было взято 10 вторых хромосом из популяции *D. pseudoobscura*. В гомозиготном состоянии эти хромосомы распределяются в ряд от нормальных до субвитальных. У потомков самок, гетерозиготных по всевозможным парам этих хромосом, выделили 450 хромосом и определили выживаемость гомозигот по ним. Оказалось, что из 450 хромосом 19 были летальными и 57 полуметальными в гомозиготном состоянии. Если этот эксперимент корректен, то из него следует, что всего 10 хромосом могут породить громадную генетическую изменчивость. Однако в последующих работах, проведенных другими генетиками, факт существования многочисленных синтетических леталей не подтвердился [179, 525, 546]. Добржанский сообщил также, что, как было обнаружено его группой, некоторые летали, вопреки утверждению Мёллера, являются гетерозисными. Однако в широкомасштабном исследовании Мукаи и Ямагути [380] никаких данных в пользу этого положения не получено. Авторы пришли также к заключению, что «супервитальных» хромосом, по крайней мере у *Drosophila melanogaster*, не су-

шествует; этот вывод находится в противоречии с соответствующим утверждением Добржанского. Кроме того, было показано, что отсутствие постулировавшейся Добржанским корреляции между приспособленностью гетерозиготы и суммой приспособленностей составляющих ее хромосом в гомозиготном состоянии обусловлено объединением леталей и умеренно вредных хромосом, имеющих разные степени доминантности [381, 580]. К тому же, как впервые отметил Стёртевант (см. [60, 70]), в природных популяциях частоты леталей слишком низки, чтобы обеспечить появление сколько-нибудь значительного числа сверхдоминантных или даже полностью рецессивных особей.

Несмотря на столь неполную аргументацию, эта работа Добржанского оказала огромное влияние на последующее развитие популяционной генетики, особенно в США. Во второй половине 60-х годов, когда у многих видов с помощью электрофореза был обнаружен полиморфизм ферментов, некоторые авторы считали даже, что Добржанский предсказал существование высокой степени полиморфизма в своей балансовой гипотезе. На самом же деле большая часть экспериментальных данных, которые, как думал Добржанский, подтверждают его балансовую гипотезу, оказалась неверной (см., например, [380]).

Мёллер [389] решительно выступал против положения о преобладающей роли сверхдоминирования в поддержании генетической изменчивости, не говоря уже о тезисе, сформулированном Лернером [299], который считал, что гетерозиготность *per se*, как правило, полезна. Он особенно опасался, что такие представления послужат доводом для преуменьшения вредности генетических последствий радиации. И все-таки утверждение, что Мёллер придерживался мнения о генетической однородности видов, ошибочно. В работе под названием «Интеграция, наметившаяся на симпозиуме по генетике, палеонтологии и теории эволюции» [387], он писал:

«Кроме того, следует помнить, что в большинстве природных кроссбредных популяций накопилось множество мутаций, которые возникли на протяжении сотен предшествующих поколений и еще не закрепились в популяции, но по которым она тем не менее гетерогенна. Таким образом, для популяции может быть характерна высокая наследственная изменчивость того или иного признака, позволяющая ей гибко реагировать на отбор практически в любом направлении, опять же вопреки тому факту, что мутация отдельного гена — событие чрезвычайно редкое».

Тезис о том, что главной причиной внутривидовой генетической изменчивости является балансирующий отбор, а не повторные мутации, содержится также в концепции генетического полиморфизма, предложенной Е. Фордом в 1940 г. (цит. по

[120]). Он определяет генетический полиморфизм как совместное существование в одном и том же местообитании двух или большего числа дискретных форм вида в таких соотношениях, что наиболее редкая из них не может поддерживаться только за счет повторных мутаций. Это определение, на мой взгляд, довольно своеобразно, так как в нем заранее подразумевается наличие некоего механизма поддержания полиморфизма и автоматически исключается возможность поддержания полиморфизма посредством баланса между мутационным процессом и отбором. Приведенное определение, с тех пор как оно было предложено, широко цитируется в учебниках по генетике. Форд утверждал, что чаще всего генетический полиморфизм поддерживается за счет преимущества гетерозигот (см. [120], с. 63). Доводы Форда в пользу широкого распространения сверхдоминирования по существу очень просты, как это ясно из формулировки, содержащейся в другой его книге под названием «Экологическая генетика» [119]. Он писал, что «преимущество гетерозигот в ходе эволюции приводит к возникновению полиморфизма» и что гетерозигота «имеет преимущество как перед менее, так и перед более благоприятствуемыми гомозиготами». Одновременно он говорил об очень интенсивном отборе, вопреки традиционному представлению об адаптивной эволюции как постепенном накоплении небольших мутационных изменений. Взгляды Форда разделяли его многочисленные сторонники, особенно в Англии.

Мне кажется, все они упустили из виду то обстоятельство, что методы экологической генетики позволяют обнаружить только случаи с очень высокими коэффициентами отбора. И в самом деле, выявить отбор малой интенсивности крайне трудно, что видно из экспериментальной работы Мукаи и Ямагути [380], исследовавших относительные выживаемости *Drosophila melanogaster*, гетерозиготных по летальным хромосомам. Эти исследователи взяли 691 вторую хромосому (274 хромосомы с летальными и 417 не содержащих леталей) из популяции Роли (шт. Северная Каролина). Они провели 688 скрещиваний и получили 241 гетерозиготу, свободную от леталей (+/+'), 350 гетерозигот по одной летали (+/l) и 97 гетерозигот по двум леталям (l/l'). В этом эксперименте среднее число подсчитанных мух, приходящихся на одно скрещивание, было равно примерно 1400, а сумма по всем скрещиваниям составила 968 820 (Мукаи, личное сообщение). Таким образом, для установления факта снижения выживаемости гетерозигот по рецессивным летальным мутациям приблизительно на 1,2% в среднем им пришлось пересчитать в контролируемых условиях около миллиона мух.

Наряду с концепцией, связывающей приспособленность со сверхдоминированием и эпистатическим взаимодействием, развивалось еще одно направление. Одним из его лидеров был Эрнст

Майр. Согласно Майру [351], классическая популяционная генетика, излагаемая на математическом языке частот генов, основана на упрощающем допущении о постоянной абсолютной селективной величине для каждого отдельного локуса, что, по его мнению, нелогично. Приспособленность гена скорее следует считать относительной; она зависит от таких факторов, как влияние других генов, с которыми он объединен, эволюционная история популяции, степень инбридинга, численность популяции и т. д. Как отмечал Майр, это новое направление, развивавшееся в течение ряда лет, является, так сказать, «теорией относительности в области популяционной генетики».

Подобным же образом Уоддингтон [562] в своей книге «Стратегия гена» критиковал работы Фишера, Райта и Холдейна за то, что разработанная ими математическая теория лишена двух элементов, которые должны быть присущи любой математической теории. Во-первых, она не привела к каким-либо заслуживающим внимания количественным выводам относительно эволюции, поскольку соответствующие формулы содержат такие параметры, как селективное преимущество, эффективная численность популяции, скорость мутирования и т. д., точные значения которых неизвестны. Во-вторых, она не выявляет новых зависимостей и процессов, которые могли бы объяснить явления, непонятные ранее. В своей книге Уоддингтон ввел множество неологизмов, таких, как метаотбор (по аналогии с метафизикой), гомеостатический и квазигомеостатический отбор, канализирующий отбор, противопоставляемый нормализующему отбору, непрерывный дрейф и прерывистый дрейф, ложный и истинный профиль приспособленностей, креод и т. д., которые при рассмотрении эволюционных процессов в природных популяциях оказываются большей частью бесполезными.

В 1959 г. во всем мире отмечалась сотая годовщина с момента выхода в свет работы Дарвина «Происхождение видов». Симпозиум в Колд-Спринг-Харборе, состоявшийся в этом году, проходил под названием «Генетика и дарвинизм XX века». В работе под заглавием «Где мы находимся?», открывающей сборник трудов симпозиума, Майр [352] опять говорил о новой популяционной генетике как о генетической «теории относительности». Обсуждая труды Фишера, Райта и Холдейна, он заметил: «В чем же, если будет позволено задать столь дерзкий вопрос, заключается конкретный вклад этой математической школы в эволюционную теорию?» Однако его собственная новая популяционная генетика носит исключительно вербальный характер и лишена какой-либо количественной основы, являясь в этом смысле полной противоположностью теории относительности в физике.

Несмотря на все попытки возвести синтетическую теорию эволюции на некий пьедестал, фактический прогресс, достигнутый

в то время, был очень незначителен. В целом это был период застоя, когда преобладала риторическая аргументация, а не строгий научный анализ.

К началу 60-х годов, судя по всему, было достигнуто согласие относительно того, что каждый признак формируется в процессе адаптивной эволюции с помощью естественного отбора и что среди мутантных генов почти нет селективно нейтральных. Например, Майр в книге «Зоологический вид и эволюция» ([353], см. гл. 8) утверждал, что, с его точки зрения, крайне маловероятно, чтобы какой-либо ген оставался селективно нейтральным сколько-нибудь продолжительное время. Он говорил также, что селективную нейтральность можно исключить почти а priori всякий раз, когда в природных популяциях обнаруживаются клины по полиморфным или количественным признакам. Он полагал даже, что отказ от введения такого фактора эволюции, как «генетический дрейф», помог бы внести ясность в дискуссии по эволюционной проблематике.

Кроме того, Форд в своей книге «Экологическая генетика» [119] выдвинул положение, согласно которому нейтральные гены не только очень редки, но и не могут достичь значительных частот, так как их нейтральность утрачивается в результате изменений окружающей среды и генетической конституции организма. Обоснование этого положения Форд находил в идеях Фишера [107], который всегда придерживался селекционистских взглядов. В одной из своих работ Фишер [108] писал: «Эволюция представляет собой процесс прогрессивной адаптации и ничего больше. Различия, выявляемые систематиками, — лишь вторичный побочный продукт, появляющийся случайно в процессе повышения приспособленности».

Концепция, привлекавшая для объяснения даже самых больших различий понятие положительного отбора — ее можно назвать «панселекционизмом», — была парадигмой, очень плодотворной для того времени; многие биологи придерживаются ее и сегодня. Я и сам испытал когда-то ее громадное влияние.

Панселекционизм, конечно, не нов. История эволюционного учения, особенно на рубеже нашего века, изобилует весьма привлекательными гипотезами относительно селективного преимущества тех или иных признаков. Например, один из авторов рассматривал розовый цвет фламинго как защитную окраску на фоне заката. Однако панселекционизм синтетической теории отличается гораздо большей утонченностью. В 50-е годы казалось, что все факты и эксперименты свидетельствуют в его пользу.

Согласно другому представлению, господствовавшему в то время, скорость и направление эволюции определяются главным образом положительным отбором, а мутации играют только второстепенную и незначительную роль. Существовало мнение, что даже если бы мутации полностью отсутствовали, аутбред-

ные виды, как правило, обладали бы достаточной изменчивостью из-за наличия рекомбинации и эволюционный процесс все равно продолжался бы длительное время. Первым эту идею высказал Фишер (см. с. 96 в [107]). Приверженцем тезиса о том, что темпы эволюции определяются главным образом факторами отбора, а не скоростью мутирования, был также и Мёллер [387]. Однако следует отметить, что Фишер и Мёллер имели в виду адаптивную эволюцию.

Оглядываясь назад, я не могу не думать об одной странности человеческой природы: если о каком-то научном воззрении постоянно говорят, что у него очень много сторонников, если его пропагандируют в своих публикациях высокие авторитеты, если о нем идет речь в учебных аудиториях, то постепенно формируется своего рода вера, в конце концов перерастающая в некую догму. Во всяком случае, это было время, когда позиции «панселекционистов» или «неодарвинистов» были самыми прочными за всю историю биологии — это была пора расцвета традиционной «синтетической теории» эволюции.

В соответствии с духом времени были высказаны многочисленные соображения относительно организации генных пулов и их преобразований в процессе эволюции. Однако эти выводы основывались на результатах исследований на фенотипическом уровне, и фактически никакого способа, позволяющего выяснить, что же происходит в ходе эволюции на уровне тонкой структуры гена, тогда не существовало. Как уже отмечалось, большое значение в определении приспособленности придавалось эпистатическим и гетерозисным взаимодействиям генов. Были введены (а некоторыми и приняты) такие термины, как *интегрированный коадаптированный комплекс*, *генетическая революция*, *связность* (cohesion) *генного пула*, *генетический гомеостаз*. Однако, по моему мнению, эти словосочетания являются скорее разными выражениями, чем научными терминами.

На протяжении всего этого периода, но главным образом независимо от модных веяний постепенно происходило значительное усложнение математического аппарата популяционной генетики — явление для биологии довольно необычное. Особенно примечательна теоретическая конструкция, называемая диффузионными моделями [228], в которых используются дифференциальные уравнения в частных производных, известные в математике как диффузионные уравнения. Такие модели позволяют рассматривать поведение мутантных аллелей с учетом как стохастических изменений, обусловленных случайной выборкой гамет, так и детерминированных изменений, связанных с мутациями и отбором. Этот мощный метод основан на результатах, полученных в замечательных работах Фишера и Райта. Моя собственная научная деятельность была посвящена главным образом разработке диффузионных моделей, и этому в значи-

тельной мере способствовали пионерские работы Райта. Метод диффузионных уравнений, несмотря на его приближенный характер, дает ответы на многие важные, но трудные вопросы, которые нельзя получить другими методами. Например, он позволяет определить «вероятность того, что единственный мутантный ген с произвольной доминантностью, появившийся в популяции ограниченной численности и обладающий определенным селективным преимуществом, в конечном счете распространится по всей популяции» [223, 227]. В популяционной генетике эту вероятность обычно называют *вероятностью фиксации гена*.

Несмотря на столь большие возможности, этот метод не мог найти применения в изучении проблем эволюции до тех пор, пока не наступила эра молекулярной биологии. Это связано с тем, что теория популяционной генетики строится на основе концепции генных частот, т. е. концентраций различных аллельных генов в популяции. Традиционные же эволюционные исследования проводились на фенотипическом уровне, и никакого прямого способа установления однозначной связи между этими двумя уровнями тогда не существовало.

Все эти проблемы были решены с развитием молекулярной генетики, которое ознаменовалось двумя достижениями. Во-первых, с помощью сравнительного анализа аминокислотных последовательностей белков родственных организмов (например, сравнения гемоглобинов позвоночных) и с привлечением палеонтологических данных появилась возможность оценить темпы накопления замен аминокислот в ходе эволюции [622] и в свою очередь скорость накопления замен нуклеотидов в генах. Во-вторых, развитие электрофоретических методов сделало возможным быстрое обнаружение изменчивости ферментов в выборках особей; эти исследования выявили у многих видов настоящее изобилие полиморфных вариантов ферментов [169, 304]. На этой основе удалось оценить внутривидовую генетическую изменчивость, причем с гораздо большей точностью, чем при использовании традиционных маркеров генов.

Таким образом, наконец пришло время применить математическую теорию популяционной генетики для изучения вопроса о том, как эволюционируют гены. Тогда казалось, что последующие исследования выявят действие дарвиновского положительного отбора или выживания наиболее приспособленных и на этом, самом фундаментальном, уровне организации живого.

Для читателей, не знакомых с молекулярной генетикой, я кратко изложу основные сведения о природе генов и мутаций на молекулярном уровне. Ген, а точнее цистрон, можно считать линейным сообщением, записанным с помощью четырех букв — оснований ДНК, обычно обозначаемых А, Т, G и С, где А — аденин, Т — тимин, G — гуанин, С — цитозин. Когда ген функционирует, это сообщение транскрибируется (переписывается) на

язык другой молекулы, называемой информационной РНК и состоящей из четырех оснований, идентичных тем, из которых состоит ДНК, за исключением тимина (Т); его заменяет урацил (U). После этого информация, содержащаяся в информационной РНК, транслируется (переводится) на язык полипептидной последовательности, образующейся путем соединения двадцати разных аминокислот. И наконец, полипептидная цепь определенным образом складывается, превращаясь в функциональную молекулу белка. Замечательно, что таким путем образуются все белковые молекулы — ферменты (например, цитохром *c*), гормоны (такие, как инсулин) и структурные белки (например, кератин), — без которых немислимо существование всего живого. Генетическое сообщение, о котором говорилось выше, составлено так, что набор из трех последовательных букв, или триплет, образует кодовое слово, или кодон, отвечающее определенной аминокислоте. Так как в каждом положении

Таблица 2.1. Стандартный код РНК

1	2				3
	U	C	A	G	
U	Phe	Ser	Tyr	Cys	U
U	Phe	Ser	Tyr	Cys	C
U	Leu	Ser	Term.	Term.	A
U	Leu	Ser	Term.	Trp	G
C	Leu	Pro	His	Arg	U
C	Leu	Pro	His	Arg	C
C	Leu	Pro	Gln	Arg	A
C	Leu	Pro	Gln	Arg	G
A	Ile	Thr	Asn	Ser	U
A	Ile	Thr	Asn	Ser	C
A	Ile	Thr	Lys	Arg	A
A	Met	Thr	Lys	Arg	G
G	Val	Ala	Asp	Gly	U
G	Val	Ala	Asp	Gly	C
G	Val	Ala	Glu	Gly	A
G	Val	Ala	Glu	Gly	G

Ala — аланин (A); Arg — аргинин (R); Asn — аспарагин (N); Asp — аспарагиновая кислота (D); Cys — цистеин (C); Gln — глутамин (Q); Glu — глутаминовая кислота (E); Gly — глицин (G); His — гистидин (H); Ile — изолейцин (I); Leu — лейцин (L); Lys — лизин (K); Met — метионин (M); Phe — фенилаланин (F); Pro — пролин (P); Ser — серин (S); Thr — треонин (T); Trp — триптофан (W); Tyr — тирозин (Y); Val — валин (V); Term. — кодон, терминирующий синтез белковой цепи. Буквы в скобках — однобуквенные обозначения аминокислотных остатков, введенные Дэйхоф.

триплета может находиться одна из четырех букв, то число кодонов равно 4^3 , или 64. 61 из них используется для кодирования 20 аминокислот, а три остальных служат знаками препинания («терминации цепи»). Все 64 кодовых слова уже расшифрованы (табл. 2.1). Мутации генов представляют собой изменения в сообщении на уровне ДНК и могут быть разделены на две группы: 1) замены одного нуклеотида на другой и 2) структурные перестройки, включающие делеции и вставки одного или большего числа нуклеотидов, а также транслокации и инверсии более протяженных участков ДНК. Известно, что при сравнении гомологичных генов близких видов замены нуклеотидов обнаруживаются чаще, чем структурные перестройки. Другими словами, самый распространенный тип эволюционных изменений на молекулярном уровне — замены нуклеотидов.

2.2. Появление мутационно-дрейфовой гипотезы (гипотезы нейтральности)

Картина эволюционных преобразований, проясняющаяся сейчас в ходе молекулярно-генетических исследований, по моему мнению, совершенно не соответствует той, которая вырисовывалась с позиций неodarвинизма. К примечательным особенностям молекулярной эволюции, обнаруженным вскоре после начала этих исследований, следует отнести примерное постоянство скорости аминокислотных замен (т. е. числа таких замен в год) для любого белка в различных родословных, кажущуюся случайность этих замен и очень высокую общую скорость изменений ДНК, равную по крайней мере одной мутационной нуклеотидной замене на геном за каждые два года, если провести экстраполяцию на содержание ДНК в гаплоидном геноме млекопитающих.

Эта скорость была получена на основании заниженной оценки, в соответствии с которой аминокислотная замена в полипептидной цепи из 100 аминокислотных остатков происходит один раз за $28 \cdot 10^6$ лет, а также в предположении, что геном млекопитающего содержит около $4 \cdot 10^9$ пар нуклеотидов. Тогда же я обратил внимание, что для млекопитающих такая оценка скорости мутационных замен на геном на поколение по крайней мере в несколько сот раз выше хорошо известной оценки, полученной ранее Холдейном [163]. Его оценка была основана на концепции «цены естественного отбора», которая позднее была названа мною «грузом замещений» (substitutional load)¹ [224]. Холдейн рассматривал ситуацию, когда исходный ген терял свое преимущество в результате изменения условий окружающей среды, а мутантный аллель с низкой приспособленностью ста-

¹ В других главах наряду с этим используется термин «субституционный груз». — *Прим. ред.*

новился более приспособленным. Затем он оценил смертность, вызванную отбором, или «цену», которую вид должен платить за сохранение мутантного гена. Холдейн показал, что цена D , которую необходимо заплатить за фиксацию с помощью естественного отбора одной генной замены, не зависит от коэффициента отбора, но зависит от исходной частоты замещающего мутантного гена. Он получил изящный результат, состоящий в том, что в случае полудоминантной мутации $D = -2\ln p$, где D — суммарная селективная смертность, p — исходная частота мутантного гена. Например, если исходная частота благоприятного мутантного гена равна $p = 10^{-6}$, то $D = 27,6$. Если мутантный аллель полностью рецессивен, то D будет намного больше. Холдейн принял $D = 30$ — величине, характерной для реального эволюционного процесса. Если отбор оказывает слабое действие на совокупность локусов, в которых за n поколений происходит в среднем одна мутационная замена, то приспособленность вида упадет в $30/n$ раз ниже оптимума, и, следовательно, интенсивность отбора будет равна $I = 30/n$. Он предполагал, что для горотелической эволюции (эволюции со стандартной скоростью) характерно значение $I = 0,1$, и считал, что скорость замещения составляет около $1/300$. Иными словами, на каждые 300 поколений приходится одна генная замена. Он также полагал, что этот результат согласуется с фактом медленной эволюции на фенотипическом уровне.

Работа Холдейна основана на детерминистической модели, в которой не учитывается влияние случайной выборки гамет в популяции ограниченной численности. Чтобы снять это ограничение, я использовал метод диффузионных уравнений и получил значение «цены», или груза, замещений для такой популяции. В случае полудоминантных мутантных аллелей этот груз для одной генной замены равен примерно $L(p) = -2\ln p + 2$, где можно принять $p = 1/(2N)$, если иметь в виду «молекулярные мутации» (подробности см. в работе [250]). Такая приближенная формула пригодна тогда, когда мутантный аллель имеет определенное селективное преимущество перед «старым» аллелем. «Цена», подсчитанная по этой формуле, примерно на 2 больше той, которая получена на основании формулы Холдейна, т. е. разность между двумя оценками обычно относительно мала.

Я применил эту и другие аналогичные формулы в предположении, что большинство мутационных замен на молекулярном уровне фиксируются с помощью положительного естественного отбора; при таком допущении груз замещений, обременяющий каждое поколение, оказался настолько большим, что его не смог бы вынести ни один вид млекопитающих. Например, для вида, насчитывающего полмиллиона особей, $L(p) = 29,6$ при $p = 1/(2N) = 10^{-6}$, так что, даже если приравнять один год к одно-

му поколению, груз, приходящийся на поколение, будет равен примерно 30. Это означает, что для поддержания постоянной численности популяции при одновременном сохранении мутационных замен, скорость появления которых равна одной замене за два года (т. е. средний груз замещений равен 15), каждый родитель должен оставить $e^{15} \approx 3,27 \cdot 10^6$ потомков с тем, чтобы один из них выжил и стал размножаться.

Это и был основной аргумент, который я использовал, предлагая мутационно-дрейфовую гипотезу, или гипотезу нейтральной молекулярной эволюции [229]. Другой аргумент, тогда мною не приводившийся, но подразумевавшийся, состоит в том, что для объяснения высокой скорости молекулярной эволюции, протекающей под действием естественного отбора, необходимо принять допущение о необычайно большой скорости появления благоприятных мутантных генов (особенно если эти мутантные гены обладают только небольшим селективным преимуществом s_1 , например равным или меньшим 0,001). Чтобы убедиться в этом, рассмотрим вид крупных млекопитающих с продолжительностью поколения, равной шести годам. Эффективная численность популяций такого вида (не обязательно совпадающая с их фактической численностью) может быть равной или меньшей 10^5 . Используя формулу для скорости эволюции, выраженной через число мутационных замен [см. уравнение (3.22) в гл. 3], я рассуждал следующим образом: поскольку $v = k/(4N_e s_1)$, где v — скорость мутирования в расчете на гамету, k — скорость появления мутационных замен в расчете на поколение, N_e — эффективная численность популяции, то, приняв $k=3$ (что эквивалентно скорости замещений, равной 0,5 замен в год), $2N_e=10^5$ и $s_1=10^{-3}$, мы получим $v=0,015$, или 1,5% на гамету и на поколение. Это очень высокая скорость возникновения благоприятных мутантных генов, поскольку приведенная величина имеет тот же порядок, что и общая скорость появления летальных и полуметальных мутаций в расчете на гамету. Очевидно, однако, что мутантные гены, обладающие преимуществом, должны появляться гораздо реже, чем вредные мутантные гены. При больших k или меньших N_e и s_1 v станет еще выше. Предположение о том, что благоприятные мутантные гены возникали с такой высокой и постоянной скоростью, по моему мнению, противоречит основному принципу адаптивной эволюции путем естественного отбора. Приняв это предположение, следует ожидать, что каждая замена благоприятного мутантного гена должна, как правило, понижать вероятность последующих мутаций, приводящих к возникновению еще более благоприятного гена, если не происходит каких-либо коренных изменений окружающей среды.

Переходя теперь к проблеме внутривидовой генетической изменчивости, необходимо отметить, что в этой области также

были получены выдающиеся результаты. Как мы уже говорили, применение методов электрофореза привело начиная с 1966 г. к неожиданному обнаружению громадной генетической изменчивости и показало, что полиморфизм на молекулярном уровне охватывает значительную часть локусов, причем у самых разных видов. Согласно Левонтину и Хабби [304], для 18 случайно взятых локусов *D. pseudoobscura* средняя гетерозиготность, приходящаяся на локус, равна около 12%, а доля полиморфных локусов составляет 30%. Сходные результаты, приведенные в сообщении Харриса [169], были получены и для человека. Таким образом, у этих видов должны существовать тысячи локусов с высокой частотой расщепления и любая особь может быть гетерозиготна по сотням или даже тысячам локусов. Однако такой полиморфизм не связан с видимыми фенотипическими отличиями и обычно не коррелирует с условиями окружающей среды. Если бы эти варианты поддерживались главным образом с помощью сверхдоминирования по отдельным локусам, то в случае вышних животных суммарный сегрегационный груз [59] или сопутствующая селективная смертность были бы, как мне представляется, недопустимо большими.

Сегрегационный груз создается вследствие появления вредных гомозигот, которые выщепляются в каждом поколении. Рассмотрим локус с парой сверхдоминантных аллелей A_1 и A_2 ; пусть относительные приспособленности генотипов A_1A_1 , A_1A_2 и A_2A_2 равны соответственно $1-s_1$, 1 и $1-s_2$ ($0 < s_1 \leq 1$, $0 < s_2 \leq 1$). В этом случае средняя приспособленность особи из равновесной популяции будет меньше приспособленности гетерозиготы на величину

$$L_s = s_1 s_2 / (s_1 + s_2). \quad (2.1)$$

Если $s_1 = s_2 = s$, то $L_s = s/2$. Так, например, если коэффициент отбора, направленного против гомозигот, равен 1%, то сегрегационный груз составит $L_s = 0,005$. Предположим, что имеется n сверхдоминантных локусов с независимым расщеплением и мультипликативным вкладом в приспособленность; в этом случае «доля селективной элиминации», связанной с расщеплением по n локусам, в расчете на особь составит $1 - e^{-sn/2}$. Приняв $s = 0,01$ и $n = 2000$, мы получим $1 - e^{-10} \approx 0,9999546$. Это означает, что если расщепление происходит по 2000 сверхдоминантных локусов с 1%-ным преимуществом гетерозигот по каждому из них, а отбор осуществляется в результате гибели менее приспособленных гомозигот, то каждая особь должна в среднем давать приблизительно 22 000 потомков, чтобы численность популяции из поколения в поколение поддерживалась на постоянном уровне. Совершенно очевидно, что ни один вид млекопитающих не может позволить себе такого расточительного воспроизводства.

В 1967 г. в ходе анализа этих интересных, но неожиданных результатов у меня возникла мысль, что большинство нуклеотидных замен, происходящих в процессе эволюции, должно быть следствием случайной фиксации селективно нейтральных или почти нейтральных мутаций, а не положительного дарвиновского отбора и что многие системы ферментативного полиморфизма селективно нейтральны и поддерживаются посредством баланса между возникновением мутаций и их случайной элиминацией. Эту гипотезу я сформулировал на заседании «Генетического клуба» в городе Фукуока, состоявшемся в ноябре 1967 г., а также в кратком сообщении, отправленном в «Nature» в декабре 1967 г. и опубликованном в феврале следующего года [229].

Позднее, в 1969 г., моя гипотеза получила мощную поддержку (особенно в той ее части, которая касается молекулярной эволюции) в статье «Недарвиновская эволюция», написанной Джеком Лестером Кингом и Томасом Джуксом. Эти авторы независимо пришли к той же идее, что и я, подкрепив свою концепцию множеством убедительных данных, почерпнутых из молекулярной биологии. Единственное различие заключалось в том, что в то время как Кинг и Джукс [268] почти не рассматривали полиморфизма белков, моя теория нейтральности касается как молекулярной эволюции, так и в равной мере механизма поддержания белкового полиморфизма [229, 230].

Предположение о том, что аллели, ответственные за полиморфизм ферментов, не подвергаются жесткому отбору и могут поддерживаться с помощью мутационного давления или очень небольшого преимущества гетерозигот, было выдвинуто также Кроу [61]. Годом раньше аналогичное мнение о почти нейтральном характере этого полиморфизма в современных популяциях было высказано Робертсоном [488] в связи с данными Харриса [169] о полиморфизме белков.

Еще раньше, до того как были описаны многочисленные системы ферментативного полиморфизма, Кимура и Кроу [246] теоретически изучали вопрос о числе изоаллелей, которые могут поддерживаться в популяции ограниченной численности. В этой работе мы рассматривали три возможных случая: 1) систему селективно нейтральных изоаллелей, частота которых в популяции определяется скоростью мутирования и случайным дрейфом; 2) систему взаимно гетерозисных аллелей; 3) смесь гетерозисных и вредных мутантных аллелей. Анализируя сегрегационный груз конечной популяции, мы пришли к выводу, что сегрегирующие локусы, характеризующиеся сверхдоминированием по приспособленности, вероятно, составляют меньшинство (вопреки утверждению Уоллеса [563]). Поэтому мы думали, что типичная дрозофила будет, по всей видимости, гомозиготной по большинству своих прошедших отбор генов. Однако это положение ни в коей мере не исключает возможно-

сти существования высокой гетерозиготности, связанной с селективно нейтральными изоаллелями. В статье мы так и написали: «Нельзя исключить возможности существования в очень больших популяциях множества практически нейтральных, высокомутабельных изоаллелей, несмотря на отсутствие экспериментальных данных, свидетельствующих об их существовании». Такие экспериментальные данные были получены спустя два года после опубликования этой работы.

Как уже отмечалось, простое объяснение факта существования многочисленных систем белкового полиморфизма в рамках модели нейтральных изоаллелей было дано Кимура [229, 230]. Несколькими годами позже Кимура и Ота [256] разработали более обоснованную и доказательную теорию. В этой работе мы выдвинули положение, что эволюционные замены аминокислот и существующий полиморфизм белков — не два независимых феномена, а просто два разных аспекта одного и того же явления, вызванного случайным дрейфом нейтральных или почти нейтральных аллелей в популяциях ограниченной численности. Короче говоря, полиморфизм белков — это одна из фаз молекулярной эволюции.

Появление теории нейтральной, или недарвиновской, эволюции вызвало дискуссию, известную под названием «дискуссии между нейтралистами и селекционистами». Доводы за и против теории нейтральности обсуждались на различных научных форумах и во многих публикациях (см., например, *Proceedings of the Sixth Berkeley Symposium on Mathematical Statistics and Probability volume V: Darwinian, Neo-Darwinian, and Non-Darwinian Evolution*, ed. L. Le Cam, J. Neyman, and E. L. Scott (1972), University of California Press; *Trends in Biochemical Sciences* 1: N152—N154).

Аргументы в пользу «нейтрального» и «селективного» механизмов обсуждаются в статьях Кроу [64, 66] и Харриса [170] и в книгах Колдера [36], Левонтина [302] и Руффи [496, 497]. Особую известность получила широко цитируемая книга Левонтина, содержащая очень подробный обзор и критическую оценку как селекционистской, так и нейтралистской позиций. Левонтин утверждает, что название «теория нейтральных мутаций» неадекватно, и предлагает более подходящий, с его точки зрения, термин — «неоклассическая теория», имея в виду дилемму классической и балансовой теорий, сформулированную Добржанским [84]. Классическая теория в интерпретации Добржанского по существу представляет собой карикатуру на взгляды Мёллера и сформулирована лишь для того, чтобы тут же ее опровергнуть, поэтому терминология, связанная с ней, не может быть адекватной. К тому же изложение теории нейтральности, приведенное в книге Левонтина, до некоторой степени односторонне, и я хотел бы внести здесь необходимые коррективы.

Мой тезис о том, что наблюдаемая скорость молекулярной эволюции слишком велика, чтобы ее можно было объяснить действием положительного отбора в рамках концепции цены отбора, сформулированной Холдейном, немедленно подвергся острой критике в работе Мэйнарда Смита [348]. Он показал, что если принять «пороговую модель», в которой определенная доля, скажем 50%, особей, имеющих благоприятные аллели в количестве, превышающем пороговое, доживает до репродуктивного возраста, то число генных замен, фиксирующихся без появления тяжелого генетического груза, может быть гораздо больше, чем в случае независимого действия отбора на разные локусы. Например, когда в каждом поколении отбирается 50% популяции, то число локусов, по которым возможен одновременный отбор с интенсивностью 0,01, при пороговом отборе достигает 25 500, тогда как в модели мультипликативной приспособленности оно равно только 138 (в обоих случаях рассматривается популяция гаплоидов; частота наиболее благоприятного аллеля каждого локуса равна 0,5). Он доказывал, что предположение о межлокусной мультипликативности приспособленностей, которое лежит в основе холдейновской концепции цены, справедливо только в исключительных случаях и, следовательно, вывод Холдейна о том, что для замены одного гена обычно требуется 300 поколений, неверен. Развивая свою концепцию, Мэйнард Смит утверждал, что скорость эволюции должна быть на порядок или даже на несколько порядков выше; по его мнению, мой вывод о том, что значительная доля генных замен на молекулярном уровне селективно нейтральна и закрепились в процессе дрейфа, хотя и может быть верным, не следует с неизбежностью из приведенных мной соображений относительно цены отбора.

Думаю, что некоторые допущения в модели Мэйнарда Смита нереалистичны. Во-первых, он принимает, что приспособленность особей определяется исключительно генами, т. е. считает наследуемость равной 100%. Однако вполне возможно, что наследуемость (в широком смысле) приспособленности очень низка. Мукаи, Шафер и Кокерхэм [376] в опытах на *Drosophila melanogaster* показали, что приспособленность, рассматриваемая как количественный признак, действительно характеризуется крайне низкой наследуемостью в широком смысле (H^2), равной примерно 0,004, и, таким образом, пороговый, или отсекающий, отбор не может быть эффективным без очень жесткой выбраковки, т. е. его работа оплачивается дорогой ценой. Во-вторых, Мэйнард Смит неявно предполагает, что благоприятные мутантные гены существуют в каждом из громадного множества локусов в любой момент времени и для выявления их преимущества не нужно ждать, пока произойдут какие-либо изменения окружающей среды. Такое предположение кажется

мне весьма нереалистичным. Например, если бы оно было верным, эволюция протекала бы гораздо быстрее у видов с очень большой численностью популяций, чем у видов с популяциями небольшого размера. Ничего подобного на молекулярном уровне не наблюдается. Скорость молекулярной эволюции в расчете на год для индивидуальных белков, таких, например, как α -цепь гемоглобина, приблизительно постоянна в самых разных родословных (см. гл. 4). Я думаю, гораздо реалистичнее предположить, как это сделал Холдейн [163], что благоприятные мутантные гены возникают из имеющихся в запасе вредных мутантных генов в результате изменений окружающей среды. Модель, сходная с моделью Мэйнарда Смита, была предложена Сведом [537], однако этот автор осознавал, что лимитирующим фактором в модели такого типа является наличие благоприятных мутаций. Расчеты Мэйнарда Смита подвергались критическому анализу (правда, с иных позиций, чем здесь) и были признаны нереалистичными в работе О'Доналда [427].

Мысль о том, что в рамках модели порогового, или отсекающего, отбора снимается также проблема обременительного сегрегационного груза (даже в случае полиморфизма по очень большому числу локусов, поддерживаемого с помощью сверхдоминирования), ранее высказывалась Кингом [267], Сведом и др. [539] и Милкмэном [356]. Модели, предложенные этими авторами, в сущности очень похожи, так что я остановлюсь только на модели Кинга. В ней предполагается, что каждый гомозиготный локус вносит вклад, равный одной единице, в величину, называемую «параметром выживаемости», который зависит также от факторов окружающей среды. Далее предполагается, что особи, у которых значения этого параметра больше определенной пороговой величины, не выживают. Другими словами, допускается, что отсекающий отбор имеет место не только при селекции животных, но и в природе. Используя эту модель, Кинг показал, что большое число гетерозисных локусов, каждый из которых обладает значительным селективным преимуществом (равным, скажем, 0,01), может поддерживаться с помощью умеренной селективной элиминации (составляющей, скажем, 50%). Одна из трудностей, с которыми сталкивается эта модель, состоит в том, что она, как отметил Левонтин [302], прогнозирует слишком большое понижение приспособленности при инбридинге. Рассмотрим, например, один из случаев, приведенных в работе Кинга [267]. Приняв среднюю долю выживающих особей за 50% и компоненту дисперсии, обусловленную сверхдоминированием, равной 0,033, можно убедиться в возможности поддержания 1000 локусов, для каждого из которых частота гетерозигот равна 0,3, а селективное преимущество гетерозигот над любой из гомозигот составляет 1%.

Если мы рассчитаем соответствующие величины для полностью гомозиготных геномов (т. е. для случая, когда инбридинг равен 100%), то окажется, что выживаемость этих гомозигот меньше одной десятичной (10⁻⁴). Это противоречит результатам экспериментальных исследований на дрозофиле [546].

Хотя в цитируемой работе оценивалась лишь жизнеспособность, существуют экспериментальные данные, свидетельствующие о том, что различия по приспособленности примерно в пять раз больше различий по одной только жизнеспособности; это также противоречит выводу о громадной инбредной депрессии, вытекающему из оценок Кинга.

Тем не менее эти модели, основанные на концепции отсекающего отбора, охотно используются некоторыми селекционистами как более приемлемая для них альтернатива теории нейтральности при объяснении молекулярной эволюции и молекулярного полиморфизма. Однако при этом, как правило, не предпринимается никаких попыток получить доказательства, что отбор этого типа действительно существует и на самом деле служит причиной замещения генов и поддержания генетической изменчивости в природе.

Я считаю, что за десять лет, прошедших после появления теории нейтральности, количество данных в ее пользу увеличилось, особенно в результате последних исследований, показавших, что нуклеотидные замены, не приводящие к замещениям аминокислот (так называемые синонимичные, или молчащие, замены), в ходе эволюции фиксируются с гораздо более высокими скоростями, чем несинонимичные замены [422, 501].

Поскольку естественный отбор действует через фенотипы организмов, в формировании которых определяющую роль играют структура и функция белков, следует ожидать, что мутации, не приводящие к замене аминокислот в белках, при прочих равных условиях подвергаются давлению естественного отбора в гораздо меньшей степени, чем мутации, приводящие к таким заменам. Кроме того, в последние годы появились данные о том, что самыми распространенными эволюционными изменениями на молекулярном уровне являются синонимичные замены, а также нуклеотидные замены в некодирующих участках ДНК [202, 236].

И наконец, совсем недавно было обнаружено, что некоторые гены эукариот (организмов, в клетках которых имеется оформленное ядро) содержат участки, называемые *вставными последовательностями* или *интронами*, которые не включаются в зрелую информационную РНК и поэтому не участвуют в синтезе белка (см. обзор [57]). Получены данные об очень высокой скорости нуклеотидных замен в ходе эволюции в интронах [556]. Ниже будет показано, что *в эволюции чаще всего проис-*

ходят такие изменения молекул, которые в меньшей степени подвергаются давлению естественного отбора.

Мне представляется, что Кинг и Джукс [268] проявили замечательную проницательность, высказав в 1969 г. такую мысль: «Если эволюционная дивергенция ДНК сопровождается случайной фиксацией нейтральных мутаций, то нуклеотиды в третьем положении триплета должны замещаться с большей скоростью»; при этом они приводили довод, что большинство нуклеотидных замен в третьем положении кодона синонимично и поэтому скорее всего нейтрально.

Тогда же в одной из своих работ, посвященных теории нейтральности, я писал [230]: «Недавнее открытие вырожденности кода ДНК, т. е. существования двух или большего числа триплетов оснований, кодирующих одну и ту же аминокислоту, по-видимому, свидетельствует о том, что нейтральные мутации не так редки, как считалось раньше». Однако в той же самой работе я из соображений предосторожности добавил: «Важно отметить, что, вероятно, не все синонимичные мутации нейтральны, даже если большинство из них и близки к таковым». Это замечание было вполне уместным в свете обнаруженного в то время факта, что синонимичные кодоны часто используются «не случайным» образом или в неодинаковой степени (подробное обсуждение этого вопроса см. в разд. 7.5).

***Гипотеза случайного дрейфа
селективно нейтральных мутаций
как эволюционная парадигма***

3.1. Случайные факторы, действующие на селективно эквивалентные аллели

Гипотеза случайного дрейфа селективно нейтральных мутаций (гипотеза нейтральной молекулярной эволюции или, кратко, гипотеза нейтральности) утверждает, что на молекулярном уровне эволюционные изменения и внутривидовая генетическая изменчивость обусловлены в основном не дарвиновским отбором, а случайным дрейфом мутантных аллелей, которые являются селективно нейтральными или почти нейтральными. Главный тезис теории нейтральности состоит не в том, что мутации на молекулярном уровне селективно нейтральны в строгом смысле этого слова, а в том, что их судьба определяется главным образом случайным генетическим дрейфом. Другими словами, интенсивность отбора в процессе молекулярной эволюции настолько мала, что основными ее факторами следует считать мутационное давление и случайный дрейф.

Чтобы полнее уяснить смысл теории нейтральности и ее отличия от традиционной синтетической теории, рассмотрим эволюционный процесс, в ходе которого в пределах данного вида одни мутантные гены последовательно замещаются другими. Каждая такая замена происходит в результате ряда следующих друг за другом событий, так что недавно возникшая редкая мутантная форма, обычно представленная единственной особью, в конечном счете распространяется по всей популяции и фиксируется, т. е. достигает частоты 100 %. На рис. 3.1 изображены траектории изменений частот мутантных аллелей после их появления в популяции ограниченной численности.

Несмотря на то что в каждом поколении любой популяции не слишком большой численности появляется значительное число мутантных форм, большинство из них теряется вследствие случайных процессов в течение немногих поколений [107, 254]. Часто не сознают, что это положение справедливо не только в отношении вредных и селективно нейтральных мутантных генов, но и в случае благоприятных мутантных генов, если их преимущество не очень велико. Например, если мутантный аллель обладает селективным преимуществом в 1 %, вероятность

того, что он в конце концов распространится по всей популяции, составляет только около 2% [107, 154]. В остальных 98% случаев популяция потеряет этот аллель в результате случайных процессов, не использовав его в эволюции. Таким образом, существует громадный разрыв между общим числом благоприятных мутантных аллелей, время от времени появляющихся в ходе эволюции данного вида, и числом мутантных аллелей, действительно включившихся в его генофонд (фиксировавшихся). Мутантные гены, которым удается увеличить свои частоты и распространиться в пределах данного вида, достигнув в конце

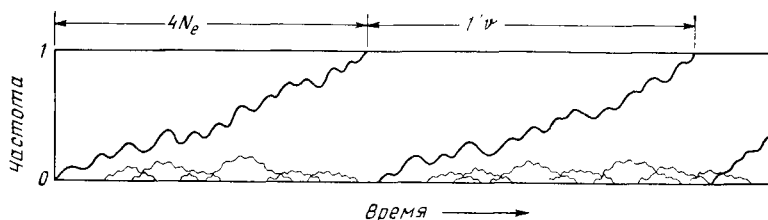


Рис. 3.1. Судьба мутантных генов после их появления в популяции ограниченной численности. «Траектории» изменения частот мутантных аллелей, достигших фиксации, изображены жирными линиями. N_e — эффективная численность популяции, v — скорость мутирования.

концов состояния фиксации, составляют лишь удачливое меньшинство. Каждое такое событие обычно реализуется на протяжении большого числа поколений. Для фиксации селективно нейтральных мутантных аллелей требуется в среднем $4N_e$ поколений, если исключить случаи, когда эти аллели теряются [253]. На рис. 3.1 траектории изменений частот мутантных генов, достигших фиксации, изображены жирными линиями.

Согласно традиционной точке зрения, мутантные гены должны обладать некоторым селективным преимуществом, чтобы закрепиться в популяции данного вида. При этом лишь изредка селективно нейтральные аллели могут перемещаться на «попутном транспорте» (т. е. за счет сцепления с благоприятным геном), достигая таким образом высокой частоты, но это уже побочное явление. С другой стороны, теория нейтральности основана на хорошо известном в популяционной генетике факте, что селективное преимущество мутантных генов не является непременным условием распространения некоторых из них по всей популяции. Если мутантные аллели селективно эквивалентны тем, из которых они возникли, то их судьба зависит от случая и соответствующие частоты будут с течением времени увеличиваться или уменьшаться случайным образом. Такие флуктуации обусловлены не только малочисленностью исходных мутантных генов, но и ограниченностью размеров популяций; в каждом поко-

лении в формировании особей следующего поколения участвует относительно небольшое число гамет из громадного их количества, продуцируемого самками и самцами. Хотя подавляющее большинство таких нейтральных мутантных аллелей теряется случайным образом, их ничтожное меньшинство в конечном счете зафиксировается. Если нейтральные мутации на молекулярном уровне широко распространены и если такой случайный процесс протекает непрерывно в течение весьма длительного времени, скажем на протяжении десятков миллионов поколений, то генетический состав популяции изменится очень сильно. Даже если вероятность фиксации мутантного гена равна всего лишь 10^{-6} , в популяции, достаточно большой для того, чтобы обеспечить ежегодное появление 10 мутантных генов на локус, будет происходить 10 мутационных замен на локус за каждый миллион лет. В популяционной генетике этот процесс известен под названием «случайного генетического дрейфа» (или, кратко, «случайного дрейфа»).

Здесь я хотел бы подчеркнуть, что мы должны ясно понимать разницу между *генной мутацией на уровне особи и мутационной (генной) заменой на популяционном уровне*. В литературе по молекулярной эволюции, говоря о различиях между гомологичными белками по определенным аминокислотам, не так уж редко имеют в виду просто «случайно распределенные точковые мутации», однако такое выражение по существу затушевывает разницу между двумя данными понятиями; аминокислотные различия между видами — это *результат* мутационных замен, а не сами генные мутации.

3.2. Случайный генетический дрейф, обусловленный ограниченной численностью популяции

Поскольку концепция случайного генетического дрейфа является основой теории нейтральности, я остановлюсь на ней более подробно. Под случайным генетическим дрейфом понимается случайная флуктуация генных частот в популяции, вызванная случайной выборкой гамет в процессе воспроизведения популяции. Общее число особей любого вида, размножающегося половым путем, не только ограничено, но, кроме того, может рассматриваться как результат случайной выборки, извлеченной из гораздо большей совокупности (или «генного пула») мужских и женских гамет, образуемых родительскими поколениями. «Размах» флуктуаций генных частот (т. е. концентраций различных аллелей) будет тем больше, чем меньше популяция.

Рис. 3.2 иллюстрирует изменение генных частот в модельной популяции, состоящей из четырех особей, при переходе от одного поколения к другому. На этом рисунке два аллеля, A_1 и A_2 , изображены черными и белыми кружками; за одно поколение

частота «черного» аллеля случайным образом изменяется от $4/8$ (т. е. от 50%) до $3/8$ (т. е. до 37,5%). И вообще: если в данный момент времени частота аллеля A_1 в популяции равна x , то среднее изменение ее за одно поколение будет равно нулю, а

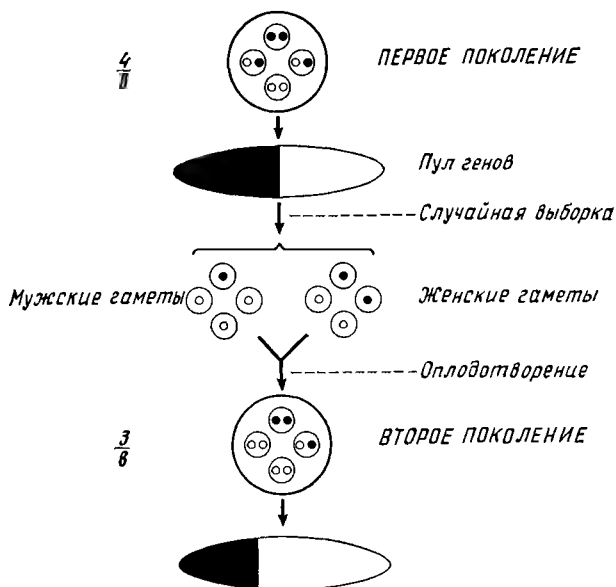


Рис. 3.2. Иллюстрация процесса случайного изменения генных частот на протяжении одного поколения в популяции, состоящей из четырех размножающихся особей. (Соответствует популяции из четырех гермафродитных особей, при воспроизведении которых происходит случайное объединение гамет.)

дисперсия (т. е. среднее квадратичное отклонение) — $x(1-x)/(2N)$, где N — численность популяции в следующем поколении. Процесс случайного дрейфа можно проиллюстрировать также с помощью рис. 3.3, на котором приведено несколько кривых изменения генных частот («траекторий выборок»), начинающихся с частоты, равной 50%.

Существует и другой метод представления, удобный для математической интерпретации случайного дрейфа. В этом случае процесс описывают с помощью распределений вероятностей генных частот. Рассмотрим ограниченную совокупность популяций (из N скрещивающихся особей каждая), в которых происходит расщепление по паре аллелей A_1 и A_2 . Обозначим частоту аллеля A_1 через x ; тогда частота A_2 будет равна $1-x$. Предположим, что в нулевом поколении ($t=0$) все популяции этого множества характеризуются одинаковой частотой A_1 (равной, скажем, p), и рассмотрим изменение долей популяций, имеющих

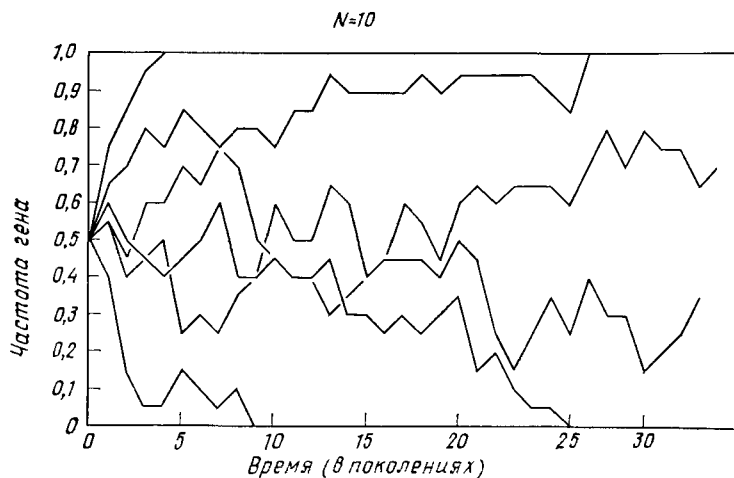


Рис. 3.3. Несколько «траекторий», отражающих обусловленный случайной выборкой гамет процесс дрейфа генных частот от начальной частоты, равной 50%. Эти траектории взяты из набора траекторий, полученных в ходе машинных экспериментов с применением метода Монте-Карло, в которых моделировалась популяция, состоящая из десяти размножающихся особей ($N=10$), с начальной частотой гена, равной 0,5.

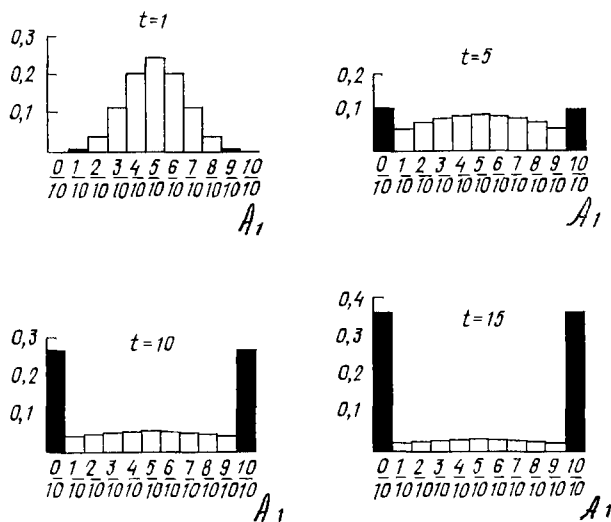


Рис. 3.4. Гистограммы, отображающие изменение распределения генной частоты во времени (приведены соответствующие распределения для 1, 5, 10 и 15-го поколений) в результате случайного дрейфа, обусловленного процессом выборки гамет. Популяция состоит из пяти гермафродитных особей, скрещивающихся случайным образом ($N=5$), и характеризуется начальной частотой A_1 (при $t=0$), равной 50% (т. е. $p=0,5$). По оси абсцисс отложена частота A_1 в популяции. Черные столбцы слева отвечают доле популяций, в которых аллель A_1 потерян, а черные столбцы справа — доле популяций, где A_1 зафиксирован.

разные частоты A_1 (т. е. изменение распределения x), в последующих поколениях.

На рис. 3.4 представлено изменение соответствующего распределения при $N=5$ и $p=0,5$. По оси абсцисс отложена частота A_1 , а высота столбцов равна доле популяций, имеющих определенную частоту A_1 . Два крайних частотных класса ($x=0$ и $x=1$) изображены черными столбцами.

Как видно из этого рисунка, распределение частот классов с незафиксированными аллелями с течением времени становится все более плоским и в конце концов A_1 или фиксируется в популяции, или теряется ею (т. е. фиксируется A_2). После 10 поколений это распределение становится почти плоским; уплощение происходит со скоростью 0,1 за поколение, т. е. со скоростью, равной $1/(2N)$.

При большей численности популяции, скажем при $N \geq 100$, гистограммы, представляющие распределение генных частот, можно с достаточной степенью точности аппроксимировать гладкой частотной кривой. Кроме того, в этом случае временной параметр t может считаться непрерывным ($t \geq 0$), хотя время по-прежнему измеряется в единицах, равных одному поколению.

Пусть $\phi(p, x, t)$ — плотность вероятности того, что частота аллеля A_1 в t -м поколении станет равной x , если при $t=0$ она равна p . Тогда можно показать, что $\phi(p, x, t)$ удовлетворяет диффузионному уравнению

$$\frac{\partial \phi}{\partial t} = \frac{1}{4N} \frac{\partial^2}{\partial x^2} [x(1-x)\phi], \quad (0 < x < 1), \quad (3.1)$$

где $\phi = \phi(p, x, t)$. С выводом этого уравнения читатели могут ознакомиться в работе Кимуры [228] и книге Кроу и Кимуры ([67]; см. с. 382—383). Точное решение этого уравнения, удовлетворяющее условию, что исходная частота гена равна p , было получено Кимурой [222] и имеет следующий вид:

$$\begin{aligned} \phi(p, x; t) = & \sum_{i=1}^{\infty} p(1-p) i(i+1)(2i+1) F(1-i, i+2, 2, p) \times \\ & \times F(1-i, i+2, 2, x) e^{-i(i+1)t/4N}, \end{aligned} \quad (3.2)$$

где $F(., ., .)$ — гипергеометрическая функция.

Рис. 3.5 иллюстрирует изменение распределения частоты гена для тех популяций, в которых не произошло ни фиксации, ни утраты этого гена, в двух следующих случаях: исходная частота p равна 0,5 (рис. 3.5, А) и 0,1 (рис. 3.5, Б). Из рис. 3.5, А видно, что если $p=0,5$, кривая распределения становится почти плоской после $2N$ поколений и при этом гены остаются незафиксированными приблизительно в 50% случаев. С другой стороны, при $p=0,1$ (рис. 3.5, Б) для получения практически плоской кривой требуется $4N$ или $5N$ поколений. Можно показать,

что и в общем случае после смены большого числа поколений распределение популяций с частотой гена $0 < x < 1$ становится плоским и что скорость уплощения равна $1/(2N)$ за поколение.

Можно подумать, что такое плоское распределение не представляет интереса, так как за время его формирования большинство генов или фиксируется, или теряется. Почему же тогда Фишер и Райт придавали ему такое значение? Это объясняет-

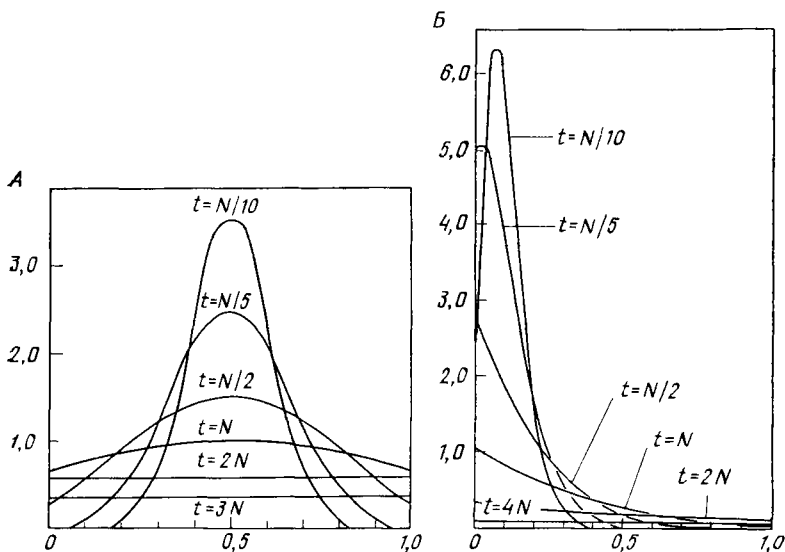


Рис. 3.5. Изменение частотного распределения аллеля A_1 под влиянием случайного генетического дрейфа. На рис. А исходная частота p аллеля A_1 равна 0,5, а на рис. Б — 0,1. t — время, а N — эффективная численность популяции. По оси абсцисс отложена частота x аллеля A_1 , а по оси ординат — плотность вероятности ϕ [222].

ся тем, что они рассматривали не группу популяций, каждая из которых «стартует» с одной и той же частоты, а группу популяций, начинающих с широкого множества исходных частот.

3.3. Эффективный размер популяции

В реальных популяциях структура скрещивания, по-видимому, гораздо сложнее, чем в модельных, рассмотренных в предыдущем разделе. Поэтому желательно свести такие сложные ситуации к эквивалентным им простым случаям, математический анализ которых был бы намного проще. С этой целью Райт [586] ввел понятие эффективной численности (или размера) популяции, которое стало одним из основных при рассмотрении проблемы случайного генетического дрейфа.

Здесь я кратко подытожу некоторые важные результаты, полученные в этой области (более подробное изложение см. в гл. 3 книги Кимуры и Оты [257]). Далее через N_e мы будем обозначать эффективную численность (размер) популяции, а через N — ее фактическую численность.

1. Если популяция состоит из N_m самцов и N_f самок, участвующих в размножении, эффективный размер популяции будет равен [586]

$$N_e = \frac{4N_m N_f}{N_m + N_f}. \quad (3.3)$$

Из этой формулы видно, что, если число размножающихся самцов и самок сильно различается, эффективный размер зависит главным образом от числа особей менее многочисленного пола. В крайнем случае, когда число самцов пренебрежимо мало по сравнению с числом самок, $N_e = 4N_m$. Таким образом, если некая популяционная группа возглавляется одним самцом, ее эффективный размер будет равен всего четырем даже при наличии бесконечно большого числа самок.

2. Если N родителей передают следующему поколению разное число гамет (k) и если численность популяции постоянна ($\bar{k}=2$), а скрещивание происходит случайным образом, то

$$N_e = \frac{4N - 2}{\sigma_k^2 + 2}, \quad (3.4)$$

где σ_k^2 — дисперсия k [592]. Отметим, что, если число потомков подчиняется распределению Пуассона, так что $\sigma_k^2 = \bar{k} = 2$, эффективный размер будет приблизительно равен фактической численности. Если же дисперсия больше среднего значения — соотношение, наиболее вероятное для большинства природных популяций, — эффективная численность будет меньше фактической. С другой стороны, если родители оставляют совершенно одинаковое число потомков ($\sigma_k^2 = 0$), мы получаем $N_e = 2N - 1$, т. е. в этом случае эффективный размер популяции примерно в два раза больше ее фактической численности. Вряд ли такая ситуация встречается в природе, однако ее можно воспроизвести в экспериментальных популяциях, контролируемых человеком.

3. Если число особей в популяции периодически изменяется, причем период изменения относительно мал и составляет n поколений, эффективный размер равен

$$N_e = n \left/ \sum_{i=1}^n (1/N_i) \right. \quad (3.5)$$

Другими словами, N_e равно гармоническому среднему (т. е. величине, обратной среднему значению обратных величин) чис-

ленностей особей за время одного периода [592]. Таким образом, эффективный размер определяется главным образом периодом малых численностей. Например, если популяция каких-либо насекомых за пять поколений увеличивается в геометрической прогрессии от 10 до 10^6 особей и затем в течение следующих пяти поколений возвращается к своему исходному размеру, $N_e \approx 54$. Периодическое понижение численности размножающихся особей может играть важную роль в значительном уменьшении эффективного размера популяции по сравнению с ее средней фактической численностью. О влиянии такого эффекта («горлышка бутылки») на генетическую изменчивость популяций читатели могут прочесть в книге Нея [401]. Эффективный размер уменьшается и в том случае, если различия особей по плодовитости характеризуются высокой наследуемостью [486, 416].

4. Если поколения перекрываются, то, как показали Кроу и Кимура [68], N_e стационарной популяции примерно равно

$$N_e = N_0 \bar{l}, \quad (3.6)$$

где N_0 — число особей, рождающихся ежегодно, τ — средний репродуктивный возраст, а

$$\bar{l} = \int_0^{\infty} l_y^2 b_y dy, \quad (3.6a)$$

где l_y — вероятность того, что особь доживает до возраста y , $b_y dy$ — ожидаемое число рождений на особь в возрасте между y и $y + dy$. Если в течение репродуктивного периода рождаемость остается постоянной, а смертность низка, формула (3.6) по существу становится идентичной формуле Нея и Имайдзуми [412]:

$$N_e = N_{\text{ср}} \tau, \quad (3.7)$$

где $N_{\text{ср}}$ — число рождающихся в течение года особей, доживающих до среднего репродуктивного возраста, τ — средний репродуктивный возраст. Значительных успехов в разработке этой проблемы достиг Хилл [180]. Он показал, что эффективный размер популяции со случайным скрещиванием и перекрывающимися поколениями равен эффективному размеру популяции с дискретными поколениями, пополняемой в каждом поколении таким же числом новых особей и имеющей ту же самую дисперсию числа потомков, порождаемых за время жизни. Им была получена формула для случая гермафродитных диплоидов:

$$N_e = (4N - 2) / (\sigma_n^2 + 2), \quad (3.8)$$

где N — число ежегодно рождающихся особей, L — продолжительность поколения в годах, σ_n^2 — дисперсия размера семьи за

время жизни одного родителя или, точнее, дисперсия (в расчете на особь) числа гамет, формирующих следующее поколение. Отметим, что многие вновь рождающиеся особи могут не оставить потомства. (Упрощенный вывод этой формулы см. в работе [180].)

5. Если данный вид состоит из большого числа более или менее изолированных субпопуляций и если их элиминация и последующее замещение происходят достаточно часто, эффек-

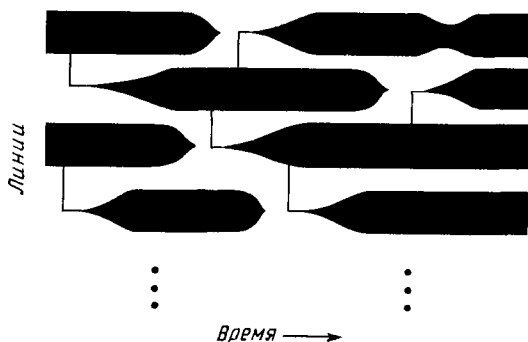


Рис. 3.6. Диаграмма, изображающая структуру популяции гаплоидного вида, размножающегося бесполом путем и состоящего из большого числа линий, которые подвергаются элиминации и последующему замещению.

тивный размер вида будет гораздо меньше, чем в случае, когда вид образует единую совокупность свободно скрещивающихся особей, или, иными словами, панмиктическую единицу. Это свойство особенно важно при рассмотрении генетической изменчивости видов бактерий (гл. 9). Можно показать (см. [342]), что в случае гаплоидного вида, состоящего из n размножающихся бесполом путем линий, каждая из которых подвергается случайной элиминации и последующему замещению (см. рис. 3.6), эффективный размер популяции данного вида будет равен

$$N_e = \tilde{N} + n/(2\lambda) + n\tilde{N}v/\lambda, \quad (3.9)$$

где $\tilde{N}$ — гармоническое среднее эффективного размера популяции в расчете на линию и на единицу времени, n — число линий (предполагаемое большим), λ — скорость элиминации линий, v — скорость мутирования (принимается модель с бесконечным числом нейтральных аллелей; при измерении этих скоростей используется одна и та же единица времени). Величину внутривидовой генетической изменчивости гаплоидного вида можно измерить с помощью «виртуальной гетерозиготности»,

которая определяется как разность между единицей и суммой квадратов аллельных частот. При равновесии, когда мутационное возникновение новых аллелей уравнивается случайной элиминацией старых, виртуальная гетерозиготность равна $2N_e v / (1 + 2N_e v)$. Отметим, что эффективный размер N_e для данного вида может быть на несколько порядков меньше его фактического размера, равного $n\bar{N}$, где $\bar{N}$ — среднее арифметическое эффективных численностей популяций в расчете на линию и на единицу времени. Можно показать (приняв островную модель структуры популяции), что эффективный размер диплоидного вида, состоящего из n колоний (субпопуляций), в котором наряду с элиминацией и образованием новых колоний происходит миграция между колониями со скоростью m на поколение, равен

$$N_e = \tilde{N} + n/[4(v + \lambda + m)] + n\tilde{N}(v + m)/(v + \lambda + m), \quad (3.10)$$

где v — скорость мутирования, λ — скорость элиминации в расчете на поколение. Если эта последняя величина много больше скорости миграции m особей, эффективный размер вида будет значительно меньше $n\tilde{N}$ — произведения числа колоний на эффективный размер отдельной колонии. При равновесии средняя гетерозиготность равна $4N_e v / (1 + 4N_e v)$, где N_e определяется формулой (3.10).

И наконец, я хотел бы привлечь внимание читателей к тому факту, что эффективный размер популяции, используемый при рассмотрении нейтральной эволюции, представляет собой эффективный размер вида в целом (или подвида, если он образует независимую репродуктивную единицу), а не локальной популяции, представляющей часть его. Это особенно важно учитывать при общем анализе проблемы структуры подразделенной популяции.

3.4. Нейтральные и почти нейтральные мутации

Для того чтобы судьба мутантных аллелей определялась преимущественно случайным дрейфом, эти аллели не обязательно должны быть строго нейтральными, т. е. совершенно эквивалентными по приспособленности. Необходимо лишь, чтобы их селективное преимущество или «вредность», измеряемые коэффициентом отбора s , не превышали величины, обратной удвоенному эффективному размеру популяции: $|s| < 1/(2N_e)$. В своей работе 1968 г., посвященной нейтральным и почти нейтральным изоаллелям [230], я писал, что мутантный ген можно определить как «почти нейтральный», если $|2N_e s|$ много меньше единицы, т. е. если $|s| \ll 1/(2N_e)$. Ли [310] считает это требование слишком жестким и вводит более адекватное, с его точки зрения, определение почти нейтрального гена: $|s| \leq 1/N_e$.

Для выяснения судьбы аллелей, близких к нейтральным, оценим вероятность фиксации мутантного аллеля в популяции. Пусть $u(p)$ — вероятность того, что мутантный аллель, исходная частота которого равна p , в конечном счете зафиксирован в популяции. Рассмотрим простой случай отбора на уровне генов, когда в единичной дозе мутантный аллель имеет селективное преимущество s , а в двойной — $2s$, т. е. случай полудоминирования

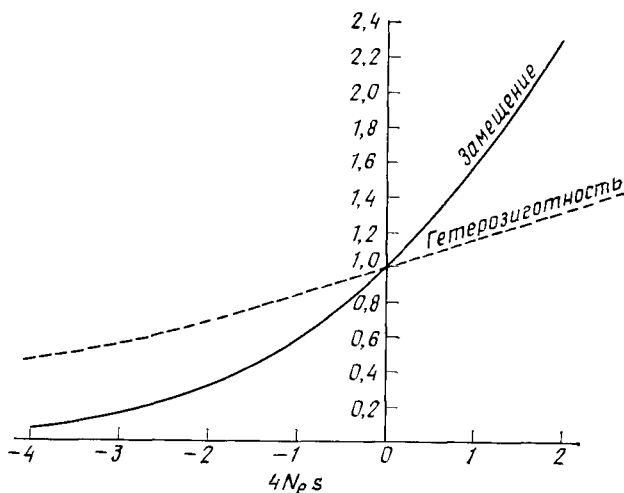


Рис. 3.7. Относительный вклад мутации, слабо подверженной отбору, в процесс замещения генов и гетерозиготность. Сплошная кривая представляет собой зависимость u/u_0 от $4N_e s$, где N_e — эффективный размер популяции, s — коэффициент отбора для мутантного гена. Пунктирная кривая — изменение $H_T/H_{T,0}$ с изменением $4N_e s$.

ния (или «отсутствия доминирования»). Можно показать [223, 227], что в этом случае

$$u(p) = (1 - e^{-Sp}) / (1 - e^{-s}), \quad (3.11)$$

где $S = 4N_e s$. Вероятность фиксации единичной мутации, обозначаемую через u , можно тогда получить, приняв $p = 1/(2N)$. Если $|s|$ мало, мы с хорошей степенью точности получаем

$$u = \frac{S}{2N(1 - e^{-S})}. \quad (3.12)$$

Коэффициент отбора (s или $S/4N_e$) в этих формулах может быть как положительным, так и отрицательным в зависимости от полезности или вредности мутантного аллеля. Когда мутация строго нейтральна ($s \rightarrow 0$), последняя формула сводится к

$$u = 1/(2N). \quad (3.13)$$

Если мы обозначим эту вероятность через u_0 и рассмотрим отношение u/u_0 , то получим

$$u/u_0 = S/(1 - e^{-S}). \quad (3.14)$$

— вклад, вносимый в осуществление генных замен мутацией, обладающей селективным преимуществом s по отношению к вкладу строго нейтральной мутации.

На рис. 3.7 соответствующая зависимость представлена сплошной кривой. Если, например, $s = -1/(4N_e)$, это отношение будет примерно равно 0,58.

Рассмотрим теперь относительный вклад мутации, мало подверженной отбору, в гетерозиготность популяции. Примем, как и прежде, допущение о полудоминантности по приспособленности. Пусть $H(p)$ — сумма долей гетерозигот, несущих мутантный аллель; суммирование производится по всем поколениям, начиная с того, в котором частота этого аллеля равна p , и кончая тем, в котором произошла его фиксация или потеря. Можно показать [232], что в этом случае

$$H(p) = 2 [u(p) - p]/s, \quad (3.15)$$

где $u(p)$ — вероятность фиксации (3.11). В случае единичной мутации $p = 1/(2N)$ и соответствующая сумма долей гетерозигот равна

$$H_T = \frac{S - 1 + e^{-S}}{Ns(1 - e^{-S})}, \quad (3.16)$$

где, как и прежде, $S = 4N_e s$. Для простоты примем предположение о равенстве фактической и эффективной численностей популяции, т. е. будем считать, что $N_e = N$. Тогда мы получим

$$H_T = 4(S - 1 + e^{-S})/[S(1 - e^{-S})]. \quad (3.17)$$

В случае строго нейтральной мутации ($s \rightarrow 0$) эта формула сводится к равенству $H_T = 2$. Введем обозначение $H_{T,0} = 2$. Таким образом, отношение

$$H_T/H_{T,0} = 2(S - 1 + e^{-S})/[S(1 - e^{-S})] \quad (3.18)$$

(где $S = 4N_e s$) представляет относительный вклад в гетерозиготность мутации, имеющей селективное преимущество s , если принять за стандарт строго нейтральную мутацию.

На рис. 3.7 соответствующая зависимость представлена штриховой кривой. Очевидно, что при отклонении s от нуля (т. е. при отклонении от строгой нейтральности) вклад мутации в гетерозиготность будет изменяться гораздо медленнее, чем ее вклад в осуществление генных замен. Например, если $s = -1/N_e$, т. е. $S = -4$, мы имеем $H_T/H_{T,0} \approx 0,46$, тогда как $u/u_0 \approx 0,075$. Это легко понять, поскольку при низких частотах вредных аллел-

лей их судьба (если только они не оказывают большого влияния на приспособленность) очень сходна с судьбой строго нейтральных аллелей из-за сильного влияния случайного дрейфа, обусловленного процессом выборки. Именно поэтому вредные аллели и могут вносить существенный вклад в гетерозиготность. С другой стороны, когда частоты вредных аллелей высоки, против них действует эффективный отбор, предотвращающий их фиксацию; по этой причине они не могут вносить значительного вклада в процесс замещения генов.

3.5. Популяционная динамика мутационной замены

Перейдем теперь к количественному рассмотрению процесса накопления новых мутаций в геноме данного вида в ходе эволюции. Рассмотрим процесс последовательного замещения мутантных генов в популяции, изображенный на рис. 3.1 жирными линиями. Возьмем полный набор нуклеотидов, составляющих геном (гаплоидный хромосомный набор). Для человека это число очень велико — порядка $3,5$ млрд. ($3,5 \cdot 10^9$). Так как скорость мутирования в расчете на нуклеотид низка (по-видимому, меньше или равна 10^{-8}), мы можем предполагать, что мутация, когда бы она ни появилась, всякий раз возникает в новом сайте, по которому в данной популяции еще не происходило расщепления мутантных форм. Это предположение [232] лежит в основе модели, известной теперь в популяционной генетике под названием «модели с бесконечным числом сайтов» (см. [234]).

Пусть v — скорость мутирования на гамету на поколение. Так как каждая особь имеет два набора хромосом, в популяции из N особей существует $2N$ хромосомных наборов, и поэтому в каждом поколении в популяции появляется $2Nv$ различных новых мутаций (для гаплоидной популяции это число равно Nv). Пусть u — вероятность фиксации единичной мутации в данной популяции. Для полудоминантных мутаций она задается выражением (3.12). Если процесс фиксации прослеживается на протяжении очень длительного времени, скорость накопления мутационных замен в популяции в расчете на поколение будет равна произведению числа новых мутаций, возникающих в каждом поколении, на вероятность их фиксации. Таким образом,

$$k = 2Nvu. \quad (3.19)$$

Отметим, что из $2Nv$ новых мутаций, появляющихся в каждом поколении, фиксации в конечном счете достигает их часть, равная u . Таким образом, k представляет скорость эволюции, выраженную через число мутационных замен.

Хотя выше мы рассматривали геном как целое, приведенные результаты можно также с хорошим приближением применить к отдельному локусу, состоящему из нескольких сотен (или

больше) нуклеотидов. Поэтому в дальнейшем мы ограничим наше рассмотрение отдельным локусом или цистроном и в выражениях для скорости эволюции будем принимать за единицу нуклеотид или кодон. Кроме того, часто оказывается более удобным рассматривать белок, кодируемый данным геном, а не сам ген.

Говоря о мутациях на молекулярном уровне, следует иметь в виду два очень важных момента: 1) число возможных аллельных состояний в любом генном локусе (цистроне) так велико, что может считаться практически бесконечным; 2) для любого короткого промежутка времени обратные (в строгом смысле) мутации столь редки, что их можно не учитывать. В качестве примера рассмотрим цистрон, кодирующий α -цепь гемоглобинов млекопитающих. Этот полипептид состоит из 141 аминокислоты и таким образом соответствует 423 нуклеотидным сайтам (на самом деле ген, кодирующий данный полипептид, содержит еще вставные последовательности, или «интроны», однако мы их не учитываем). В этом случае вследствие одних только замен оснований могут существовать 4^{423} , или примерно 10^{254} , аллельных состояний, так как в каждом сайте может находиться любое из четырех оснований. Таким образом, любой из этих аллелей может превратиться посредством одноступенчатых замен оснований в $3 \cdot 423$, или 1269, других аллелей. Вероятность возврата любого из производных аллелей к исходному состоянию в результате последующих одношаговых замен составляет — в предположении равновероятности всех замен оснований — всего $1/1269$.

Между прочим, эти соображения свидетельствуют в пользу одной частной модели аллельных мутаций (называемой «моделью с бесконечным числом аллелей»), в которой предполагается, что мутация — всякий раз, когда она возникает, — приводит к появлению нового, ранее не существовавшего аллеля [246]. Хотя эта модель пригодна для изучения уровней существующей генетической изменчивости видов (гл. 8), для анализа рассматриваемого здесь процесса мутационных замен она не подходит. В высокополиморфном локусе фиксация единичного аллеля — событие маловероятное; в данном случае мы можем наблюдать только сдвиг аллельных частот во времени. Таким образом, без идентификации мутаций на уровне индивидуальных нуклеотидных сайтов четкая интерпретация скорости мутационных замен невозможна. Именно поэтому анализ процесса фиксации мутационных замен с использованием модели с бесконечным числом аллелей, проводившийся Гэссом и Юэнсом [150], для практического применения в биологии непригоден.

Вернемся к формуле (3.19) для оценки скорости эволюции и рассмотрим два крайних случая: случай нейтральных мута-

ций и мутаций, имеющих явное преимущество. Вначале примем предположение о селективной нейтральности мутантных аллелей. В этом случае, как видно из (3.13), $u=1/(2N)$. Подставляя это выражение в формулу (3.19), мы получаем, что

$$k=v. \quad (3.20)$$

Иными словами, скорость эволюции, выраженная через число мутационных замен за поколение, независимо от размера популяции равна скорости мутирования в расчете на гамету [229]. Эта замечательная зависимость выполняется только в случае нейтральных аллелей. Однако, как показано в предыдущем разделе, если коэффициент отбора мутантного аллеля по абсолютной величине много меньше $1/(2N_e)$, его судьба сходна с судьбой нейтрального мутантного аллеля и, таким образом, указанное равенство приближенно справедливо.

Предположим теперь, что мутантный аллель имеет явное селективное преимущество; при этом $4N_e s \gg 1$. Тогда из (3.12) мы с хорошей степенью точности имеем [228]

$$u=2sN_e/N. \quad (3.21)$$

Подставив это выражение в (3.19), получим

$$k=4N_e s v. \quad (3.22)$$

Это означает, что скорость эволюции зависит от эффективного размера популяции N_e , селективного преимущества мутантного аллеля s , а также от скорости v , с которой в каждом поколении возникают благоприятные мутантные гены. В таком случае скорость эволюции должна сильно зависеть от окружающей среды, являясь высокой для видов, осваивающих новые экологические условия, и низкой для видов, обитающих в стабильной среде. Как мне кажется, крайне маловероятно, что произведение $N_e s v$ остается одним и тем же для разных линий в эволюции позвоночных — независимо от того, была ли эволюция на фенотипическом уровне очень быстрой (например, в родословной, ведущей к человеку) или же практически остановилась (например, в родословной, ведущей к карпу). Как мы увидим в следующей главе, наблюдаемая скорость эволюции для данной молекулы (скажем, молекулы гемоглобина) в расчете на год постоянна в разных линиях, а равенство (3.22) с этим фактом, очевидно, несовместимо. Равенство (3.22) не согласуется также с положением, что фактические скорости постоянны скорее в расчете на поколение, чем на год. Наблюдаемое постоянство гораздо лучше соответствует результатам, получаемым в рамках теории нейтральности [см. уравнение (3.20)], хотя разные авторы выражают в этой связи определенные сомнения, задаваясь вопросом, почему скорость возникновения нейтральных мутаций в расчете именно на год (а не на поколение) ока-

зывается одинаковой у организмов, столь сильно различающихся по продолжительности поколений. Эту проблему мы обсудим позже (разд. 8.7), при рассмотрении модели «эффективно нейтральных мутаций» [238]. Отметим, что при сравнении организмов с близкой продолжительностью поколений, но живущих в совершенно разных условиях, этой проблемы не возникает.

Анализируя популяционную динамику мутационных замен, мы должны кроме вероятности фиксации генов знать еще сред-

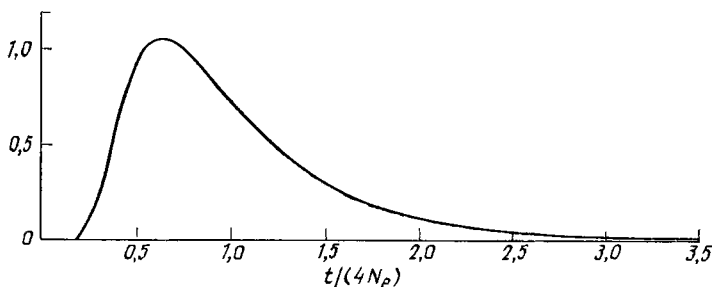


Рис. 3.8. Распределение вероятностей длин временных интервалов, прошедших до фиксации селективно нейтрального мутантного гена. По оси абсцисс отложено время в единицах, равных $4N_e$ поколениям, а по оси ординат — плотность вероятности [233].

нюю длину временного интервала, необходимого для осуществления каждой замены. Общая теория этого вопроса была разработана с применением метода диффузионных уравнений Кимуры и Отой [253]. Она позволяет оценить среднее число поколений, сменяющихся до момента фиксации (исключая случаи, когда аллели в конечном счете теряются); при этом принимается, что исходная частота мутантного аллеля равна p . В частном случае селективно нейтрального мутантного аллеля, когда $p=1/(2N)$, среднее число поколений, сменившихся до фиксации, будет примерно равно

$$\bar{t}_1 = 4N_e. \quad (3.23)$$

Другими словами, для достижения фиксации селективно нейтрального мутантного аллеля в результате случайного дрейфа требуется время, в среднем равное учетверенному эффективно-му размеру популяции. Было также показано, что стандартное отклонение для числа поколений, предшествовавших фиксации, равно около $2,15N_e$, или примерно половине средней величины [254, 397]. Кроме того, для случая нейтральных мутантных аллелей было получено распределение вероятностей длин временных интервалов, за которые происходит фиксация [233]. Это распределение изображено на рис. 3.8. Видно, что дости-

жение состояния фиксации до момента времени, равного $0,8N_e$ поколениям, — событие крайне маловероятное. Для селективно благоприятных мутантных аллелей средний временной интервал, необходимый для фиксации, будет меньше, тогда как в случае сверхдоминантных аллелей — больше.

Вообще говоря, для фиксации в популяции каждого мутантного аллеля требуется очень много времени, и поэтому, определяя скорость эволюции по формуле (3.19), мы должны рас-

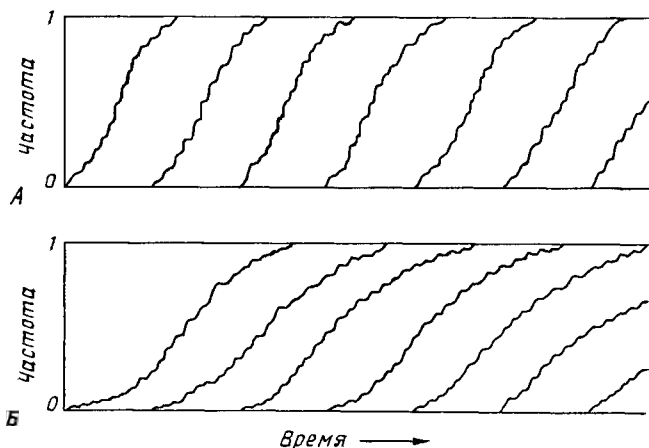


Рис. 3.9. Процессы, приводящие к мутационным заменам в популяции. Увеличение частоты индивидуальных мутантных аллелей в случае А происходит гораздо быстрее, чем в случае Б, однако скорость эволюции, выраженная через число мутационных замен, в обоих случаях одинакова.

смагивать еще большие времена. Отметим, что скорость эволюции, определяемая по формуле (3.19), не зависит от скорости увеличения или уменьшения частот отдельных мутантных аллелей в популяции. Роль играет лишь средний интервал между двумя последовательными фиксациями. Следовательно, в обоих случаях (А и Б), представленных на рис. 3.9, значения k будут одинаковыми, даже если в случае А частоты отдельных мутантных аллелей увеличиваются гораздо быстрее, чем в случае Б. Для нейтральных мутантных аллелей среднее время между двумя последовательными фиксациями равно $1/v$ поколениям, где v — скорость мутирования.

3.6. О неправильном толковании теории нейтральности и ее критике

Так как многочисленные критические высказывания и комментарии, касающиеся моей теории нейтральности, зачастую основываются на неправильном ее понимании, я хотел бы обсу-

дить здесь некоторые из них. Ни в коем случае не следует думать, будто, как ошибочно полагает Цукеркандл [621], гены, вовлеченные в эволюционный процесс, согласно теории нейтральности, лишены функциональной нагрузки. Они могут быть как свободными от функциональных ограничений, так и функционально нагруженными; при этом, однако, теория нейтральности постулирует, что мутантные формы каждого гена, участвующего в молекулярной эволюции, почти селективно нейтральны, т. е. могут одинаково успешно выполнять свою функцию, рассматриваемую с точки зрения ее роли в выживании и размножении особи.

Иногда противники теории нейтральности называют нейтральные замены эволюционным «шумом». Я, однако, думаю, что этот термин здесь не подходит. Мы же не считаем «шумом» синонимы в языке; точно так же неверно трактовать замены нейтральных аллелей просто как шум или потерю генетической информации. Тудэй [547] разделяет все пары нейтральных аллелей на четыре типа: строго нейтральные, квазинейтральные, условно нейтральные и псевдонейтральные. Он пишет, что пары строго нейтральных аллелей, очевидно, не вносят вклада в приспособленность, и если они существуют, то являются просто эволюционным шумом. По-моему, строго нейтральные аллели правильнее считать абсолютно «бесшумными».

Если нейтральные вариации представлены аминокислотными заменами в каком-либо белке, это означает, что такие замены одинаково приемлемы с точки зрения нормального функционирования белка у данного организма. Кроме того, для применимости теории нейтральности эквивалентность замен не обязательно должна быть полной. Единственное необходимое условие заключается в том, что результирующая разность приспособленностей должна быть очень небольшой, как это уже отмечалось в разд. 3.4. Левонтин в своей книге [302] (см. с. 197) пишет, что если аминокислотные замены совершенно безразличны к действию естественного отбора, то их можно называть «генетическим хламом». Очевидно, он смешивает функциональную эквивалентность с отсутствием функции. Я думаю, что «генетическим хламом» могут быть названы только генетически инертные участки генома.

Рассматривая возможность возникновения селективно эквивалентных мутаций на молекулярном уровне, я хотел бы подчеркнуть, что у организмов, особенно высших, существует физиологический гомеостаз — эффективный буфер против возмущений внутренней и внешней среды. Поэтому флуктуация условий среды никоим образом не приводит автоматически к флуктуации дарвиновской приспособленности мутантов. Этот важный момент часто упускается из виду специалистами в области математической популяционной генетики, занимающимися пробле-

мами молекулярной эволюции. Они рутинно используют термин «флуктуация среды», подразумевая при этом соответствующую ей «флуктуацию коэффициентов отбора» и не задаваясь вопросом о биологической основе характеризующихся определенной продолжительностью флуктуаций относительных приспособленностей аллелей, принадлежащих тысячам локусов (см., например, [347]).

Некоторые критические замечания в адрес теории нейтральности являются следствием ошибочного определения естественного отбора. Здесь я должен подчеркнуть, что термин «естественный отбор» должен пониматься в строго дарвиновском смысле; естественный отбор реализуется через дифференциальное выживание и воспроизведение особи. Очень часто за доказательство существования естественного отбора принимают просто наличие функциональных различий между двумя молекулярными формами, в то время как при этом необходимо исследовать еще выживаемость и плодовитость. Это часто упускают из виду биохимики, когда они принимаются за обсуждение вопросов, связанных с естественным отбором. К сожалению, термин «отбор» порой используют, искажая его смысл. Например, Цукеркандл [621] пишет, что «положительный отбор *per se* не придает биологического значения мутанту». Я думаю также, что некоторых недоразумений при трактовке теории нейтральности можно избежать, если проводить четкое различие между положительным и отрицательным отбором. Последний, приводящий к элиминации вредных мутантов, при замещении генов заметной роли не играет, за исключением случаев, когда вредный эффект мутантного аллеля настолько мал, что его подавляет случайный дрейф (разд. 3.4). Со времени появления замечательной работы Мёллера, выполненной на заре исследований генетики дрозофилы, известно, что отрицательный отбор — самая распространенная форма естественного отбора ([392]; см. также [161]). На фенотипическом уровне отрицательный отбор предстает в виде «стабилизирующего», или «центростремительного», отбора, а положительный — в виде «направленного» (подробное количественное рассмотрение этого вопроса дается в гл. 6). Существование отрицательного отбора не противоречит теории нейтральности.

Кроме того, как уже подчеркивалось, следует четко разграничивать генные мутации на уровне особи и замены мутантных форм на популяционном уровне. С молекулярной эволюцией непосредственно связаны только замены мутантных форм в популяциях. Участники очень многих дискуссий такого разграничения не проводят. Я хотел бы подчеркнуть, что, как было показано в предыдущем разделе, в случае благоприятных мутантных аллелей скорость их замен в ходе эволюции зависит не только от темпа мутирования, но также в значительной ме-

ре от численности популяции и степени селективного преимущества. Тем не менее некоторые биохимики при обсуждении проблем молекулярной эволюции принимают в расчет только скорость мутационного процесса.

Не так давно несколько авторов, в частности Цукеркандл [620] и Гиллеспи [129], заявили, что они создали «селекционистские» теории, из которых следуют те же выводы, что и из теории нейтральности. Я думаю, что все эти построения ошибочны. Рассматривая различные компоненты приспособленности в случае аминокислотной замены, Цукеркандл [620] придерживается предположения, что большинство аллелей генетически «достаточны», когда они существуют порознь, и конкурируют друг с другом при совместном существовании. Согласно этому автору, после фиксации в результате отбора селективно более благоприятного аллеля его преимущество постепенно уменьшается в результате последующих фиксаций других аллелей, что делает возможным фиксацию все новых и новых аллелей. Модель Цукеркандла трудна для понимания и количественной интерпретации, однако он, как мне кажется, полагает, что если скорость мутирования постоянна, то скорость мутационных замен, происходящих под действием естественного отбора при неизменных условиях среды, тоже постоянна. Очевидно, Цукеркандл не учитывает влияния размера популяции на накопление благоприятных мутаций и тем самым на скорость эволюции. В большой популяции число благоприятных мутаций, появляющихся в каждом поколении, намного больше, чем в малой. К тому же различные мутации необязательно будут возникать в той последовательности, которую постулирует Цукеркандл.

«Модель со случайной средой» Гиллеспи [129] предполагает флуктуацию коэффициентов отбора с определенной автокорреляцией, однако в ней совершенно не учитывается вклад мутаций. Если принять в расчет мутации, то, как показали Такахаата и Кимура [541], выводы Гиллеспи, касающиеся гетерозиготности, потребуют коренного пересмотра. Так или иначе, все эти псевдонейтралистские теории имеют одну специфическую особенность: согласно им, естественный отбор действует таким образом, что его последствия оказываются такими же, как если бы никакого отбора не существовало. Можно, наверное, предложить селекционистскую модель, имитирующую какое-то одно из следствий нейтралистской модели, однако представляется совершенно невероятным, чтобы существовала селекционистская модель, имитирующая *все* следствия теории нейтральности. С аналогичной трудностью сталкиваются математики, моделирующие детерминистические процессы получения последовательности случайных чисел; с такой же трудностью встретился бы недобросовестный экспериментатор при попытке сфаб-

риковать большое количество фиктивных данных, характер которых не смогла бы выявить ни одна комбинация статистических тестов.

Некоторые авторы, например Гэтлин [125], критикуют теорию нейтральности на том основании, что расположение нуклеотидов в геноме не является совершенно случайным. Однако, как подчеркивали Кимура и Ота [263], функциональные ограничения вполне совместимы с существованием селективно нейтральных замен. Например, если некая группа аминокислот должна быть гидрофильной, в соответствующих кодонах могут происходить случайные замены, но образующиеся при этом последовательности нуклеотидов будут неслучайными. Рассматриваемый участок, конечно, испытывает давление отрицательного отбора против гидрофобных аминокислот, однако, как отмечалось выше, существование отрицательного отбора не противоречит теории нейтральности. Теория нейтральности не постулирует эквивалентности *всех* аминокислот в определенном сайте; она только предполагает, что большинство эволюционных изменений связано с возникновением таких мутантных форм, которые являются селективно эквивалентными. Помимо функциональных ограничений следует учитывать и возможное различие скоростей мутирования для четырех разных оснований. Кроме того, как показал Кох [275], частота мутационных замен для данной пары оснований может зависеть от типа соседних пар оснований.

Хорошо известен неслучайный характер встречаемости соседних нуклеотидов в геноме позвоночных, у которых динуклеотид СрG встречается реже, чем можно было бы ожидать, исходя из нуклеотидного состава ДНК (см., например, [201]); эта неслучайность приписывается действию естественного отбора [534]. Недавно, однако, было показано, что данный феномен объясняется главным образом различием скоростей мутирования. Согласно Бёрду [20], наблюдаемый «дефицит» динуклеотидов СрG в ДНК позвоночных связан с метилированием ДНК. Наиболее распространенной метилированной последовательностью в ДНК у животных является mCpG, и для нее характерна тенденция к мутационному превращению в TpG, что приводит после репликации к появлению TpG и CpA. Это означает, что элиминация двух СрG сопряжена с появлением одного TpG и одного CpA. Бёрд [20] показал, что для ДНК животных характерна тесная корреляция между дефицитом СрG и степенью метилирования. У ДНК позвоночных степень метилирования наиболее высока и в то же время наблюдается самый острый дефицит СрG, тогда как ДНК насекомых очень слабо метилирована и не обнаруживает сколько-нибудь значительного дефицита СрG. Частично метилированная ДНК иглокожих по дефициту СрG располагается между этими двумя типами ДНК. Ис-

пользуя обширные данные относительно различных групп животных, Бёрд установил также факт существования замечательной корреляции между дефицитом SrG и избытком $\text{TrG} + \text{CrA}$. Эти результаты подтверждают предположение Салсера [501], который считает метилированные SrG мутационными «горячими точками», и находятся в противоречии с тезисом Субак-Шарпа и др. [534], согласно которому недостаток SrG в геноме позвоночных связан с действием естественного отбора.

Таким образом, у нас нет никаких доказательств того, что неслучайность, обнаруженная Гэтлин, вызвана адаптивными мутационными заменами, а не функциональными ограничениями и отрицательным отбором. Позже в качестве свидетельства против теории нейтральности приводились данные о неодинаковом (или «неслучайном») использовании синонимичных кодонов, однако такие данные тоже можно объяснить обычными ограничениями и отрицательным отбором. Более подробно проблема неслучайного использования синонимичных кодонов будет рассмотрена в разд. 7.5.

Некоторые критические выступления и скептические высказывания в адрес теории нейтральной молекулярной эволюции основаны на ошибочном мнении, будто, согласно этой теории, абсолютно все мутации на молекулярном уровне строго нейтральны. Лично я никогда так не думал, и теория нейтральности этого, конечно, не предполагает. Существование интенсивного отрицательного отбора и определенного положительного отбора не противоречит теории нейтральности, согласно которой *большинство* эволюционных замен обусловлено случайным генетическим дрейфом селективно почти эквивалентных мутантных аллелей.

Сопоставление темпов молекулярной и фенотипической эволюции

4.1. Некоторые особенности фенотипической эволюции

Формы эволюции (на примере эволюционной истории позвоночных)

Прежде чем рассматривать темпы эволюции на молекулярном уровне, я кратко остановлюсь на особенностях фенотипической эволюции, выявляемых при изучении ископаемых остатков. Это поможет мне показать, что характер молекулярной эволюции совершенно отличен от характера фенотипической эволюции.

В ходе палеонтологических исследований, выполненных за последние сто лет, было получено много ценной информации относительно эволюционной истории начиная с кембрийского периода (600—500 млн. лет назад), когда впервые появились многочисленные ископаемые остатки. Особого внимания заслуживает выяснение филогении позвоночных — группы, к которой принадлежим мы сами (см. авторитетные сводки Ромера [492, 493] и Мак-Фарланда [321]). Поскольку знакомство с эволюционной историей позвоночных понадобится нам в следующем разделе при обсуждении скоростей (темпов) эволюции глобинов, я резюмирую здесь сведения о некоторых интересных в данной связи аспектах филогении позвоночных.

Древнейшие из известных настоящих позвоночных — бесчелюстные (Agnatha), ископаемые остатки которых находят в отложениях ордовикского периода (500—440 млн. лет назад; см. геохронологическую схему на рис. 4.2). Костные пластины таких животных обнаруживаются начиная с позднего кембрия. Длина тела у них варьировала от нескольких сантиметров до полуметра, и благодаря непарным плавникам они выглядели довольно странно. Вероятно, эти животные обитали в морской воде и питались, всасывая воду с содержащейся в ней пищей посредством сильных мышечных сокращений; воду они фильтровали через жабры, расположенные по обе стороны позади ротового отверстия. Их ископаемые остатки особенно многочисленны в силурийском периоде (440—400 млн. лет назад). Тело этих животных было покрыто панцирными пластинами, и поэтому их называют остракодермами (щитковыми).

Обычно считают, что панцирь служил для защиты от гос-

подствовавших тогда хищных членистоногих — эвриптерид (или ракоскорпионов). Из костных пластин панциря возник внутренний скелет, сначала сформировавшийся вокруг головной части, а позднее развившийся в то целостное внутреннее образование из множества костей, которое мы находим у высших рыб. Без возникновения костного скелета появление более прогрессивных классов позвоночных было бы невозможно; обсуждая этот факт, Ромер [492] замечает, что мы должны быть благодарны эвриптеридам — врагам наших предков.

Новая эпоха в эволюции позвоночных началась в конце силурийского периода, когда у них развились челюсти, образовавшиеся путем увеличения одной из пар дуг, прежде поддерживавших жаберные щели. В течение девонского периода (400—350 млн. лет назад) произошли широкое расселение и адаптивная радиация первых челюстноротых. В то же время бесчелюстные остракодермы постепенно вырождались и к концу девона вымерли; до наших дней дошли лишь немногие из их потомков, представленные такими примитивными и специализированными формами, как миноги и миксины.

Следующим важным этапом эволюции позвоночных было завоевание суши, начатое амфибиями и завершенное рептилиями. Геологические данные свидетельствуют, что во время девонского периода существовало много мелких лагун и были обычны сезонные засухи. В этих условиях больше шансов выжить должны были иметь рыбы, способные с помощью своих плавников вылезать из высыхающих водоемов и переползать к другим водоемам, заполненным водой. Поэтому плавники в конце концов превратились в ноги. В условиях частых засух преимущество получали также те особи, которые могли использовать плавательный пузырь для воздушного дыхания. Группа кистеперых рыб — кроссоптеригий — дала начало предкам наземных позвоночных. Расцвет кистеперых пришелся на девонский период, а к концу палеозоя их численность резко упала, вероятно, из-за того, что они уступили в конкурентной борьбе амфибиям, которые произошли от них и обитали в той же самой среде.

Долгое время считалось, что кистеперые полностью вымерли. Однако несколько десятилетий назад в Индийском океане на большой глубине был пойман экземпляр дожившей до наших дней реликтовой рыбы, принадлежащей к целакантам и названной латимерией. Это событие было настоящей научной сенсацией, а латимерия стала, по-видимому, самым знаменитым живым ископаемым.

Первые четвероногие позвоночные (тетраподы) появились в девонском периоде. Это были амфибии, которые для размножения должны были возвращаться назад в воду, где они, как и современные лягушки, откладывали яйца и развивались. Среди ранних амфибий были переходные формы (такие, как их-

тиостега), обладавшие четырьмя короткими ногами, но, подобно рыбам, сохранившие хвостовой плавник. От амфибий произошли рептилии, которые откладывают яйца амниотического типа, защищенные твердой скорлупой. Это новое «изобретение» сделало возможным переход к полностью наземному образу жизни. Вероятно, указанное эволюционное изменение возникло, помимо прочего, как прямая адаптивная реакция на сезонные засухи: яйца, откладываемые в водоем, были обречены на гибель при его высыхании. Таким образом, откладка яиц в сухую почву, по-видимому, появилась раньше, чем произошел полный отказ взрослых особей от водного образа жизни.

Рептилии появились к каменноугольному периоду (350—270 млн. лет назад). Переселению животных из воды на сушу должны были предшествовать растения. Широкое распространение наземных сосудистых растений произошло в девонском периоде, а в следующем, каменноугольном периоде они достигли расцвета, образовав обширные леса на поверхности Земли. Это создало условия для быстрой экспансии насекомых, которая произошла в конце каменноугольного периода и обеспечила подходящую «кормовую базу» для широкой радиации рептилий, пришедшей в пермском периоде и в следующей за ним мезозойской эре.

Континенты Земли были тогда совершенно иными, чем теперь. Как впервые предположил более полувека назад немецкий геофизик Альфред Вегенер, все участки суши когда-то были объединены в единый суперконтинент, названный Пангеей. Позднее он раскололся на части, расходящиеся в процессе дрейфа, что в конечном счете привело к образованию современных континентов. Гипотеза дрейфа континентов, предложенная Вегенером, долгое время подвергавшаяся нападкам, на самом деле по существу оказалась верной, как видно из результатов, недавно полученных наукой о Земле.

Пангея сформировалась к концу палеозоя, т. е. примерно 230 млн. лет назад. К концу триасового периода (225—180 млн. лет назад) ее безраздельными хозяевами [36] стали динозавры. В течение мезозойской эры (225—70 млн. лет назад) происходила широкая радиация пресмыкающихся. Среди них были не только динозавры, но также летающие рептилии, морские рептилии и предки птиц. Хотя большинство таксонов из преобладающей группы рептилий — архозавров — к концу мезозоя вымерло, птицы весьма преуспели и выделились в особый класс позвоночных, представители которого дожили до наших дней. Их способность к полету, которая обеспечивает защиту от наземных хищников, требует больших энергетических затрат, поэтому в отсутствие хищников проявлялась тенденция к утрате этой способности и к превращению птиц в наземных обитателей.

В пермском периоде (270—225 млн. лет назад) появились зверообразные пресмыкающиеся — терапсиды, которые господствовали на суше вплоть до триаса (226—180 млн. лет назад). Затем численность их стала уменьшаться, и на протяжении большей части мезозоя они «уступали лидерство» динозаврам. По-видимому, эффективный способ передвижения на двух ногах, свойственный господствовавшим рептилиям с сильно выраженными хищническими наклонностями, был более выигрышным в конкурентной борьбе, чем четвероногая походка терапсид.

Первые млекопитающие появились в конце триаса около 200 млн. лет назад. Скучные палеонтологические данные свидетельствуют о том, что наши мезозойские предки, относящиеся к млекопитающим, были небольшими, невзрачными созданиями, размерами и внешним обликом напоминавшими мышей или крыс. По-видимому, они вели преимущественно ночной образ жизни. Австралийские яйцекладущие млекопитающие, известные под названием однопроходных, или клоачных, и представленные утконосом (*Ornithorhynchus*) и ехидной (*Echidna*), образуют побочную ветвь, отошедшую от основного ствола млекопитающих на очень ранних стадиях их эволюции (см. рис. 4.2). Они дожили до наших дней благодаря географической изоляции на ареале, где у них не было конкурентов и где не существовало хищников из числа более высокоорганизованных млекопитающих.

Для основной ветви млекопитающих мезозой был временем тяжких испытаний, имевшим, однако, большое значение для формирования характерных признаков этих животных, особенно «смышлености» и гомеотермии, которые в значительной мере усовершенствовались в ходе борьбы за существование в условиях тирании рептилий. По выражению Ромера [492], «как млекопитающие, мы должны испытывать чувство благодарности динозаврам за их непреднамеренное содействие нашему становлению».

Характерной особенностью млекопитающих является то, что они выкармливают детенышей и заботятся о них, обеспечивая безопасность потомства в период, необходимый для развития мозга, который эффективно используется млекопитающими в течение всей их жизни. Формирование более развитого мозга (устройства для переработки информации) было, вне всякого сомнения, истинной причиной успеха млекопитающих. Дискуссия о *K*- и *r*-отборе [318] как об «адаптивной стратегии», столь модная теперь в математической экологии, по-видимому, почти не отражает существа проблем, которые возникают при рассмотрении направления эволюции млекопитающих. Общее направление, доминирующее в эволюции млекопитающих, выражается, по моему мнению, в тенденции к рождению небольшого числа

потомков при эффективной заботе о них, гарантирующей их выживание в случае резких изменений среды.

К концу мезозоя все динозавры, которые до этого процветали, вымерли. Истинная причина массового вымирания крупных рептилий, обитавших на земле, в воде и в воздухе, до сих пор остается одной из основных загадок палеонтологии.

Недавно Алварес и др. [5] предложили смелую гипотезу, объясняющую это вымирание на рубеже мелового и третичного периодов около 65 млн. лет назад; основой к созданию гипотезы послужил обнаруженный ими факт резкого возрастания концентрации иридия в слое глины, отложенном как раз в указанное время. Согласно их гипотезе, в то время на Землю упал большой астероид (диаметр которого оценивается в 10 ± 4 километра) и в результате удара в стратосферу поднялось огромное количество пыли. Пыль стояла в воздухе в течение нескольких лет и широко распространилась по всей атмосфере. Наступившая вследствие этого темнота привела к сильному подавлению фотосинтеза, оборвав пищевые цепи в самом их начале. Крупные травоядные и хищные животные, включая динозавров, вымерли. За последнее время эта эффектная гипотеза, судя по всему, получила дополнительное фактическое подтверждение (см. [214]).

Вначале млекопитающие не отличались большим разнообразием; их чрезвычайно широкая адаптивная радиация началась с наступлением кайнозоя, когда они стали заполнять различные экологические ниши, которые прежде были заняты рептилиями (подразделения кайнозойской эры приведены на рис. 4.1). Из двух основных групп млекопитающих — сумчатых (Metatheria) и плацентарных (Eutheria) — первые не особенно преуспели. Поистине удивительного распространения достигли млекопитающие, принадлежащие к Eutheria, что и сделало кайнозой «эрой млекопитающих» (подробнее см. книгу Куртена [286]). Это особенно очевидно при рассмотрении палеоценовой эпохи — первых 10 млн. лет кайнозойской эры. Согласно Куртену [286], в раннем, среднем и позднем палеоцене возникло соответственно шесть, десять и четырнадцать новых семейств млекопитающих, из которых к концу палеоцена вымерло только девять семейств, входящих в последнюю группу; адаптивная радиация млекопитающих происходила беспрецедентными темпами.

В царстве растений на протяжении мелового периода (но задолго до полного вымирания динозавров) бурно эволюционировали цветковые (покрытосеменные), что привело к появлению различных семейств, широко распространенных в настоящее время. Цветковые растения наряду с млекопитающими представляют формы жизни, характерные для кайнозойской эры.

Большой интерес обычно вызывает эволюция приматов, так как именно к этому отряду относится человек (приматы включают также лемуров, низших обезьян и антропоидов). Приматы ведут преимущественно древесный образ жизни. Считается, что их общим предком было животное, сходное с тупайей (*Tupaia*), которая в настоящее время обитает в Индии. Многие признаки, характерные для приматов, сформировались как приспособления

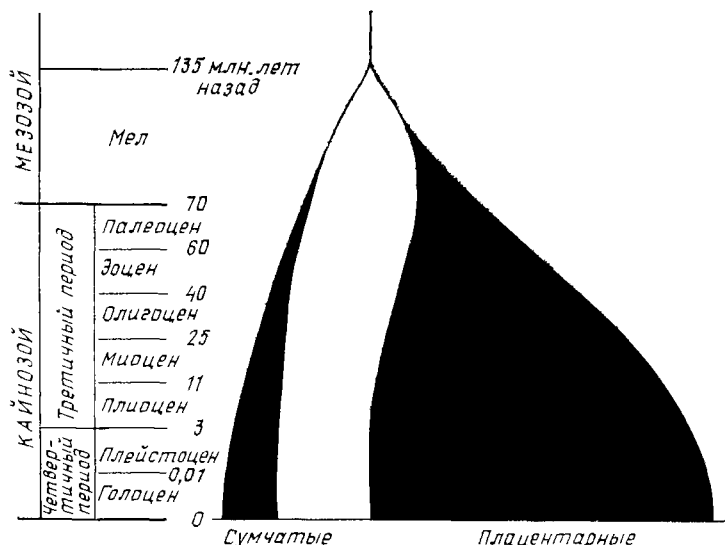


Рис. 4.1. Адаптивная радиация млекопитающих в кайнозое. Диаграмма построена по данным, приведенным в работах [286, 493].

к жизни на деревьях. Наиболее примечательные из них — руки с противопоставленными большими пальцами, обладающие мощной, цепкой хваткой, и обращенные вперед глаза, обеспечивающие хороший обзор и способность к очень точной оценке расстояния. Чтобы осуществлять управление соответствующими функциями, нужен был высокоразвитый мозг. С другой стороны, обоняние у приматов подверглось значительной редукции. Ископаемые остатки из Эфиопии свидетельствуют, что прямым предком человека был австралопитек, живший около 4 млн. лет назад.

Эволюция человека окончательно «встала на рельсы», когда около трех миллионов лет назад наши предки отказались от жизни на деревьях и перешли к прямохождению на двух ногах. Считается, что в позднетретичное время происходил процесс сокращения площади лесов и замещения их на покрытые травой саванны. Это вынудило предков человека перейти к назем-

ному образу жизни. Люди не слезали с деревьев; деревья «покинули» их.

Человек не принадлежит к числу особенно сильных созданий и, лишившись защиты деревьев, он мог выжить только благодаря своей смекалке, предпосылки которой возникли в результате развития мозга. Около 2 млн. лет назад эволюционно продвинутый гоминид с объемом мозга 750 см^3 уже изготавливал каменные орудия ([567]; датировка скорректирована). Такой мозг по объему примерно вдвое меньше, чем мозг современного человека (1400 см^3). Возможно (и это отмечал еще Мёллер [390]), между техникой применения орудий и совершенствованием генетических основ способности к их лучшему использованию существовала положительная обратная связь. Это способствовало эффективной работе естественного отбора, ведущего к быстрому развитию интеллекта.

Несмотря на обилие найденных недавно ископаемых остатков, время дивергенции человека и африканских человекообразных обезьян (шимпанзе и гориллы) остается неизвестным и является предметом многочисленных дискуссий. Согласно результатам традиционных палеонтологических исследований, основанных на сравнении анатомических структур (например, зубов), эта точка дивергенции находится где-то между 20 и 30 млн. лет назад. В частности, ископаемые остатки приматов, названных рамапитеками и сивапитеками, датируемые 14 млн. лет назад, по-видимому, свидетельствуют о ранней дивергенции человека и африканских человекообразных [511]. С другой стороны, результаты иммунологического исследования альбуминов различных видов приматов [502] указывают, что эта дивергенция произошла только около 5 млн. лет назад. Этот вывод основан на эмпирическом правиле, согласно которому логарифм иммунологического расстояния ($\log ID$) является своего рода «молекулярными часами», т. е. его увеличение во времени происходит с постоянной скоростью. Хотя никакого обоснования этому положению у авторов нет, они считают само собой разумеющимся, что молекулярные часы всегда точны, в том числе в случае родословных человека и человекообразных обезьян. Позднее Кинг и Уилсон [270] сравнили 44 ферментных локуса человека и шимпанзе и получили количественную оценку генетического различия между ними, рассчитав расстояние по формуле Нея [400]. Этот новый результат также свидетельствовал о большом генетическом сходстве человека и шимпанзе и опять привел к оценке времени дивергенции, равной примерно 5 млн. лет. К числу достижений, связанных в какой-то мере с применением молекулярных методов датирования, можно отнести установление родства орангутана с рамапитеком и сивапитеком, которых ранее считали предками человека [468] (так что никакого противоречия между датировками, основанными на сравнении инфор-

мационных макромолекул, и свидетельствами палеонтологической летописи о появлении гоминид около 4 млн. лет назад теперь не существует).

Оригинальная точка зрения на проблему времени дивергенции родословных человека и человекообразных обезьян высказывалась Кортландом [284]. Он предположил, что образование Африканской рифтовой долины и последующее ее затопление реками и озерами привело к изоляции, с одной стороны, предка гориллы и шимпанзе в Западной Африке, и с другой — предка человека в Восточной Африке. Согласно его датировке, этот разлом возник около 14 млн. лет назад, однако вполне возможно, что предложенный им механизм функционировал на десять миллионов лет позже, так как геологическая активность в рифтовой долине продолжалась на всем протяжении плиоценовой и плейстоценовой эпох. Дивергенция орангутана (ныне обитающего в лесах Азии) и африканских человекообразных обезьян, вероятно происшедшая 10 млн. лет назад, тоже может быть связана с возникновением рифта, который привел к образованию Красного моря.

Эволюционная история позвоночных, вкратце изложенная выше, ясно показывает, что определяющую роль в быстрой фенотипической эволюции играют экологические условия и что прогрессивная эволюция почти всегда представляет собой своего рода ответ организмов на вызов, брошенный средой. Происхождение человека от примитивного бесчелюстного — процесс, занявший полмиллиарда лет, — прекрасный пример прогрессивной эволюции, однако не следует забывать, что в ходе эволюции гораздо чаще происходило вырождение и вымирание. Холдейн [164] указывал, что на каждый случай прогрессивной эволюции, т. е. усложнения морфологии и поведения потомков по сравнению с соответствующими структурами и поведением их предков, вероятно, приходится десять случаев регрессивной эволюции. Главная причина, благодаря которой эволюция в целом оказалась прогрессивной, состоит в том, что вид, приобретающий новое полезное качество, давал начало различным видам-потомкам с большей вероятностью, чем вид, потерявший некоторые из таких качеств.

Одной из важнейших особенностей фенотипической эволюции является ее приспособительный характер. Наиболее наглядно эта особенность обнаруживается при рассмотрении феномена, называемого конвергенцией. Поразительный пример такого рода представляет адаптивная радиация сумчатых Австралии в сравнении с адаптивной радиацией плацентарных млекопитающих. К сумчатым относятся, в частности, сумчатый муравьед, кротобразные формы (сумчатые «кроты»), собакоподобное животное тасманийский «волк», белкообразные формы, в том числе сумчатые «летяги» и т. д. (см. [513]). Анало-

гичный параллелизм существует между североамериканскими и южноамериканскими млекопитающими (включая вымершие виды). Из этих примеров видно, что освоение сходных экологических условий животными, принадлежащими к двум разным, но родственным группам, связано с возникновением форм, ведущих сходный образ жизни. Как показывают детальные анатомические исследования, такое сходство является чисто внешним. Очевидно, что и генетические структуры, лежащие в основе соответствующих признаков представителей этих двух групп, тоже различны.

Приспособительный характер эволюции — красноречивое свидетельство в пользу того, что эволюция на уровне формы и функции определяется главным образом дарвиновским положительным отбором, который приводит к адаптации организмов в условиях окружающей их среды. Положительному естественному отбору по существу безразлично, какие генетические механизмы определяют развитие данной формы и соответствующей функции; объект его действия — целостные фенотипы. Это означает, что анализ «адаптивных стратегий» может быть полезен при рассмотрении фенотипической эволюции, но не обязательно будет иметь смысл при изучении лежащих в ее основе генетических изменений и тем более эволюционных изменений на молекулярном уровне.

Темпы фенотипической эволюции

Продолжая рассмотрение эволюции на фенотипическом уровне, мы сосредоточим теперь внимание на проблеме ее темпов (скоростей). Существование поразительных различий в темпах эволюции разных групп животных обсуждалось еще Дарвином в его книге «Происхождение видов». Он отмечал (см. гл. 10 его книги), что виды различных родов и классов изменялись с неодинаковыми скоростями или в разной степени. Например, силурийская *Lingula* лишь немного отличается от живущих ныне видов этого рода, тогда как большинство других силурийских моллюсков и все ракообразные претерпели сильные изменения. Дарвин приходит к заключению, что «наземные организмы, по видимому, изменялись с большей скоростью, чем морские». Согласно другому замечанию Дарвина, есть основания считать, что организмы, находящиеся наверху «лестницы существ», как правило, изменяются быстрее тех, которые находятся на нижних ее ступенях, хотя у этого правила есть исключения.

Позднее Ренш [481], проводивший обширные исследования возраста видов и надвидовых таксонов различных животных, пришел к выводу, что эволюция более примитивных групп морских животных протекала медленнее, чем более высокоорганизованных групп наземных животных. Этот факт он объяснял

влиянием среды, а именно тем, что морские животные в гораздо меньшей степени, чем наземные, подвержены воздействию периодических изменений таких существенных экологических факторов, как температура, влажность и т. д. Кроме того, вероятность флуктуаций численности популяций большинства морских животных меньше, чем для наземных животных; при этом давление естественного отбора, как правило, остается постоянным.

Согласно Реншу, появлению совершенно нового структурного типа часто предшествует период «взрывной» эволюции (тахителии); следующая за ним фаза медленных изменений (брадителии) может сохранять определенные виды животных на протяжении очень длительного времени. Ренш обращает внимание на интересный факт: число поколений, приходящееся на единицу времени, в общем случае не связано со скоростью эволюции. Например, эволюция слонов была относительно быстрой, тогда как эволюция мелких грызунов — медленной, хотя в последнем случае число поколений на единицу времени в 40—100 раз больше, чем в первом. (Он объясняет медленные темпы эволюции грызунов большим постоянством их среды обитания и большими размерами их репродуктивных сообществ.) Еще один пример такого рода представляет эволюция человека, которая на уровне фенотипа протекала очень быстро, несмотря на исключительно большую — по сравнению с другими животными — продолжительность поколения гоминид.

Обширное и скрупулезное исследование темпов эволюции, основанное на анализе палеонтологических данных, было проведено Симпсоном [512, 514]. Симпсон различает два типа скоростей — морфологическую и таксономическую. Морфологическая скорость отражает темпы изменения морфологического признака или комплекса признаков в определенной филогенетической линии («филуме»), поэтому ее можно считать «филогенетической скоростью». Так можно оценить, например, темпы изменения длины зубов в родословной современной лошади со времени ее предка эогиппуса, жившего в нижнем эоцене.

С другой стороны, таксономическая скорость отражает время существования и эволюционный успех определенной таксономической единицы. Эту скорость можно назвать филетической в том смысле, что она дает представление о темпах появления новых родов и замещения ими старых родов в отдельной родословной. Например, в линии, ведущей к современной лошади (*Equus*) и начинающейся с гиракотерия, произошла последовательная смена восьми родов (если не считать *Equus*, время существования которого еще не истекло). Продолжительность этого филогенеза — около 60 млн. лет. Таким образом, средняя филетическая таксономическая скорость равна 0,13 родам на 1 млн. лет (т. е. время существования одного рода составляет 7,5 млн. лет). Эта величина попадает в интервал «стандартных»

скоростей, которые Симпсон назвал «горотелическими». Другой пример — эволюция триасовых и более древних аммонитов. В этом случае таксономическая скорость равна 0,05 родам на 1 млн. лет, т. е. гораздо меньше, чем при эволюции лошади. Возможно, как отмечал Симпсон [514], оценка таксономических различий, основанная на заключении квалифицированных специалистов, дает более адекватное представление об интегральных эволюционных преобразованиях, чем любая обобщенная мера различий по морфологическим признакам.

Таксономическая скорость может также определяться на основе изменений числа или частоты встречаемостей различных таксономических групп. Скорость таких изменений можно называть частотной таксономической скоростью. Хотя прямой связи с темпами фенотипической эволюции в обычном понимании этот параметр не имеет, мне представляется уместным привести здесь соответствующий пример из монографии Симпсона [514]. Симпсон представил ряд кривых зависимости «время—частота», отражающих частоты встречаемости различных семейств копытных при их первом и последнем появлении, а также общие частоты этих семейств в период от позднего мела до плейстоцена. По этим кривым можно определить скорость изменения общей частоты, скорость возникновения (в расчете на 1 млн. лет) и скорость вымирания семейств. Анализ, проведенный Симпсоном, показывает, что скорость возникновения семейств достигает максимума приблизительно в середине указанного периода (т. е. в эоцене), тогда как наивысшие темпы вымирания характерны для второй половины этого периода.

Возвращаясь к рассмотрению темпов эволюционных изменений в отдельных линиях, приведем несколько примеров очень медленной эволюции. Особого упоминания заслуживают «живые ископаемые» («бессмертные», как называл их Симпсон [513]) — ныне живущие виды, появившиеся в очень далеком прошлом и с тех пор почти не изменившиеся. К числу наиболее известных живых ископаемых принадлежат морской моллюск *Lingula* (не отличающийся от своего ордовикского предка, жившего более 400 млн. лет назад), похожее на ящерицу пресмыкающееся *Sphenodon* (практически не изменившееся с юрского периода), опоссум (сохранившийся неизменным с мелового периода) и крокодилы (не изменившиеся со времен раннего мела). Эти животные должны были хорошо адаптироваться к окружающей их среде, не изменявшейся в течение очень длительного времени. Действие отрицательного отбора, элиминирующего мутации, было столь эффективным, что на фенотипическом уровне все время поддерживалось определенное status quo. С другой стороны, у нас нет оснований полагать, что такое сохранение status quo обусловлено снижением темпов мутирования. Предлагаемая интерпретация согласуется с фактами

эпизодического возникновения побочных ветвей (позднее вымиравших) — так, например, опоссум дал начало ветви южноамериканских хищных сумчатых.

Согласно Симпсону [513], многие «живые ископаемые» характеризуются следующими общими особенностями. Большинство этих «бессмертных» имеет относительно небольшие размеры; они образуют довольно крупные локальные популяции, обладают широкой адаптацией и живут в относительно постоянных условиях. При своем возникновении они были представлены в основном прогрессивными формами, которые впоследствии, при первой представившейся возможности, претерпели быструю эволюцию, приведшую к появлению новых адаптивных типов; они настолько хорошо приспособились к условиям окружающей среды, что формы, появившиеся позднее, уже не могли их вытеснить.

И наконец, я хотел бы упомянуть об одной работе Холдейна [160], посвященной проблеме измерения темпов эволюции («морфологических скоростей», по терминологии Симпсона [514]). По мнению Холдейна, при сравнении скоростей эволюционных преобразований различных органов или какого-то одного органа в разных родах соответствующие изменения лучше всего выражать в процентах. Если средняя длина определенной морфологической структуры (скажем, зуба) за время t увеличилась от x_1 до x_2 см, то скорость ее изменения за единицу времени можно оценить по предложенной им формуле

$$(\ln x_2 - \ln x_1)/t. \quad (4.1)$$

Согласно оценкам Холдейна, основанным на данных Симпсона [512], высота одного определенного зуба лошади увеличивалась в ходе ее эволюции в среднем примерно на $4 \cdot 10^{-8}$ в год, или на 4% за 1 млн. лет. Единица скорости эволюции, которую он предложил называть *дарвином*, представляет собой скорость изменения количественного признака, равную 10^{-6} в год; иными словами, один дарвин соответствует увеличению или уменьшению величины признака в e раз (e — основание натурального логарифма, равное 2,718...) за 1 млн. лет. Скорость эволюции лошади, выраженная в этих единицах, составляет около 40 миллидарвинов. Используя тот же метод, Холдейн показал, что длина тела динозавров, принадлежащих к шести подотрядам, на протяжении мезозойской эры увеличивалась со скоростью около $2,6 \cdot 10^{-8}$, или около 26 миллидарвинов; эта скорость несколько меньше скорости эволюции лошади. Если применить формулу Холдейна к увеличению объема мозга в процессе эволюции гоминид (обсуждавшейся в предыдущем подразделе), то мы получим величину, составляющую около 1/5 дарвина, так как объем их мозга удваивался примерно за каждые 3 млн. лет. Увеличение мозга гоминид — пример очень быстрой эволю-

ции на уровне фенотипа. Когда эти скорости превышали один дарвин, как правило, действовал такой фактор, как вмешательство человека, которое послужило причиной очень быстрых изменений домашних животных и культурных растений.

За время, прошедшее с тех пор, когда эти вопросы обсуждались Симпсоном и Холдейном, среди палеонтологов развернулась бурная дискуссия относительно характера фенотипической эволюции, которую одни считают постепенной и существенно непрерывной, а другие — прерывистой, т. е. складывающейся из внезапных «событий видообразования», сменяемых периодами эволюционного застоя. (Обзор различных мнений по этому поводу дается в сводке Шопфа [504].) Если эволюция носит прерывистый характер, то скорости изменений, происходящих в период видообразования, должны быть гораздо выше соответствующих оценок Холдейна; далее каждый вид становится, приобретая устойчивость к эволюционным преобразованиям, живым ископаемым и существует до тех пор, пока его не вытеснит другой. В то время, когда я работал над этой книгой, споры относительно темпов эволюционных изменений были в самом разгаре, и поэтому я, не будучи палеонтологом, еще не выработал свою точку зрения на предмет этой дискуссии.

4.2. Скорость эволюции на молекулярном уровне

Сравнительное изучение аминокислотных последовательностей белков

Давней мечтой специалистов по генетике популяций было определение скорости замещения генов в процессе эволюции видов. Эта величина должна более однозначно характеризовать темпы эволюции, чем любая другая мера скорости, основанная на сравнении фенотипов. Симпсон [514] высказывал мнение, что идеальным показателем темпов эволюции является «генетическая скорость», хотя тогда было неясно, как ее измерить. Этот вопрос разрешился позже, с развитием молекулярной генетики, изучающей молекулярную структуру генов. Примечательно, что впервые возможность получения достоверных оценок скоростей, с которыми новые мутантные гены включаются в геномы эволюционирующих видов, появилась после проведения сравнительного анализа аминокислотных последовательностей белков. Как отмечал Кроу [62], прямое сопоставление аминокислотных последовательностей белков позволяет обойти классический генетический анализ и проводить генетическое сравнение даже таких видов, как, например, мышь и кит, которых нельзя скрестить друг с другом.

За последние пятнадцать лет накоплено множество данных такого рода, что нашло отражение в составленной Дэйхоф [74,

75, 76, 77] обширной сводке по аминокислотной последовательности белков. Наиболее широко изучались глобины, хотя примерно такие же полные данные получены и для цитохромов *c*. Особенно активно изучаются последовательности глобинов позвоночных, которые служат прекрасной модельной системой при анализе скоростей эволюции на молекулярном уровне. Мы рассмотрим главным образом эволюцию гемоглобина и миоглобина. Гемоглобин — это белок крови, который служит переносчиком кислорода от легких к тканям и углекислого газа от тканей к легким. Функция миоглобина, существующего в форме мономера, состоит в депонировании кислорода в клетках скелетных мышц. Оба этих глобина совершенно необходимы для жизни позвоночных. У акул и высших позвоночных гемоглобин является тетрамером, состоящим из двух α - и двух β -цепей. Однако в крови морской миноги (*Petromyzon marinus*) он присутствует преимущественно в виде мономера (одиночных цепей), хотя кодируется, по-видимому, шестью разными генами [178]. Ядро молекулы гемоглобина составляет группировка, содержащая атом железа и известная под названием гема, к которой как раз и присоединяется молекула кислорода. Для более детального ознакомления со структурой и функциями белков можно обратиться к книге Дикерсона и Гейса [81], являющейся превосходным введением в этот предмет.

Сравним α -цепи гемоглобинов позвоночных. У млекопитающих α -цепи состоят из 141 аминокислотного остатка, причем α -цепи разных млекопитающих гомологичны. Так, например, если мы сравним человека и японского макака, то обнаружим, что, за исключением четырех аминокислотных сайтов, α -цепи их глобинов совершенно идентичны [346]. Таким образом, различие в аминокислотных последовательностях этих двух цепей равно всего 2,8%. Так как каждый сайт белковой последовательности может быть занят любой из 20 аминокислот, вероятность случайного совпадения двух последовательностей даже для сегмента из 10 соседних аминокислотных сайтов будет очень мала (отметим, что $1/20^{10} \approx 10^{-13}$). Из четырех различий по аминокислотам между α -цепями гемоглобина человека и японского макака три, как следует из таблицы генетического кода, можно объяснить единичными мутационными изменениями (заменами оснований) в каждом из соответствующих сайтов, а четвертое появилось в результате по крайней мере двух мутационных шагов. Очевидно, что гены, детерминирующие эти две α -цепи, произошли от общего предкового гена в результате пяти (или около этого) замен оснований. С другой стороны, α -цепи гемоглобинов человека и собаки различаются по 23 аминокислотным сайтам, что составляет 16,3% от общего числа сайтов. При сравнении еще более отдаленных таксонов различия будут еще больше. Например, α -цепи гемоглобинов чело-

века и карпа отличаются друг от друга приблизительно по 50% аминокислотных сайтов. К тому же эти две цепи различаются вставками или делециями, составляющими три аминокислот-

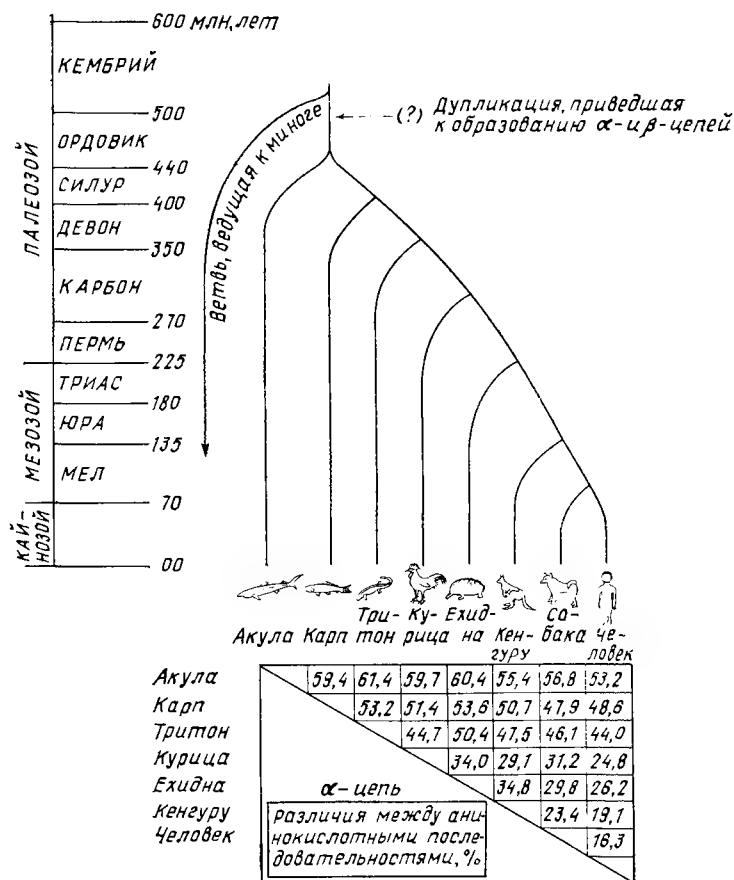


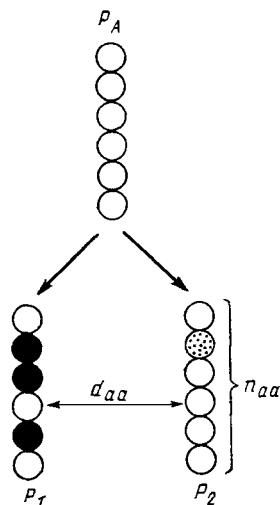
Рис. 4.2. Различия между аминокислотными последовательностями (в %) α -цепей гемоглобинов восьми позвоночных и схема филогенетических отношений, отражающая время дивергенции рассматриваемых таксонов.

ных сайта. На рис. 4.2 приведены оценки аминокислотных различий, полученные при попарных сравнениях α -цепей гемоглобинов восьми позвоночных, и схема их филогенетических отношений, отражающая время дивергенции. Таблица на этом рисунке составлена по атласу Дэйхоф ([77]; Alignment 132) и несколько отличается от приведенной Дэйхоф матрицы 34; мы

исключили из рассмотрения вставки и делеции и ограничились подсчетом аминокислотных замен.

Из рис. 4.2 можно сделать вывод о существовании четкого параллелизма между временем дивергенции и степенью аминокислотных различий. Этот параллелизм станет еще более отчетливым, если вместо простых различий в процентах рассмотреть расчетные оценки числа аминокислотных замен, действительно происшедших в ходе эволюции. Чтобы нам было проще исполь-

Рис. 4.3. Диаграмма, иллюстрирующая возникновение двух гомологичных полипептидных цепей P_1 и P_2 из одной предковой цепи P_A в результате накопления мутаций (черные кружки и кружок с точками); n_{aa} — общее число аминокислотных сайтов, d_{aa} — число сайтов, по которым различаются два полипептида. В данном иллюстративном примере $n_{aa}=6$, $d_{aa}=3$. Обратите внимание, что, хотя число зафиксировавшихся мутаций равно четырем, число сайтов, по которым можно обнаружить различия между цепями, равно только трем (случаи, когда две мутации приводят к неразличимым заменам, здесь не учитываются).



зовать эти оценки в дальнейшем, я опишу процедуру их получения в общем виде.

Пусть n_{aa} — полное число аминокислотных сайтов, по которым сравниваются две гомологичные белковые последовательности; предположим, что они отличаются друг от друга по d_{aa} сайтам (см. рис. 4.3). Доля сайтов, по которым наблюдаются различия (доля аминокислотных различий), равна $p_d = d_{aa}/n_{aa}$. Наша задача — оценить число аминокислотных замен, действительно происшедших в ходе эволюции. Если две мутации (аминокислотные замены) возникают в гомологичных сайтах в двух эволюционных линиях — одной, ведущей к P_1 , и другой, ведущей к P_2 , — или если две мутации происходят в одном и том же сайте в одной из линий, то это регистрируется как одно-единственное аминокислотное различие между P_1 и P_2 . Может случиться, что первая мутация приводит к различию по данному аминокислотному сайту, а вторая сводит это различие на нет, однако такие ситуации мы не учитываем, так как они относительно редки (составляют менее 5% от общего числа мутаций). Чтобы ввести поправку на перекрывающиеся мутации, мы

должны принять простое, но очень полезное предположение, что статистика аминокислотных замен подчиняется распределению Пуассона [622, 327].

Пусть K_{aa} — среднее число аминокислотных замен, приходящееся на пару гомологичных сайтов двух сравниваемых полипептидов. Будем считать, что вероятности 0, 1, 2, ... аминокислотных замен, происходящих в определенном сайте, задаются рядом Пуассона:

$$e^{-K_{aa}} + K_{aa}e^{-K_{aa}} + \frac{K_{aa}^2}{2!} e^{-K_{aa}} + \dots \quad (4.2)$$

Это допущение основывается на представлении, согласно которому замена в каждом сайте — событие, очень редкое для любого данного промежутка времени; однако в случае чрезвычайно длительных временных интервалов вероятность появления замены уже не будет бесконечно малой.

В частности, вероятность полного отсутствия замен по данному сайту равна $e^{-K_{aa}}$. Приравнивая это выражение доле аминокислотных сайтов, по которым два полипептида идентичны,

$$e^{-K_{aa}} = 1 - p_d, \quad (4.3)$$

мы получим

$$K_{aa} = -\ln(1 - p_d) \approx -2,3 \lg(1 - p_d), \quad (4.4)$$

где $p_d = d_{aa}/n_{aa}$ — доля аминокислотных различий. Отметим, что на рис. 4.2 приводятся значения 100 p_d . Стандартная ошибка для K_{aa} дается формулой

$$\sigma_K = \sqrt{\frac{p_d}{(1 - p_d) n_{aa}}}, \quad (4.5)$$

[231]. Строго говоря, формула (4.3) справедлива только в том случае, если скорости замен для разных сайтов равны, однако в качестве первого приближения она вполне приемлема, даже когда эти скорости различаются, если только средняя скорость замен достаточно мала (см. [407]).

Тогда скорость эволюционных замен аминокислот на сайт в год задается формулой

$$k_{aa} = K_{aa}/(2T), \quad (4.6)$$

где T — число лет, прошедших после эволюционной дивергенции двух цепей от общей для них предковой цепи. Отметим, что множитель 2 в знаменателе соответствует двум ветвям подразумеваемого филогенетического дерева.

Сравним, например, α -цепи гемоглобина человека и акулы. В 65 сайтах у них находятся одинаковые аминокислоты, а в 74 — разные. Кроме того, эти две цепи различаются вставками

или делециями (при сравнении последовательностей часто называемыми «пробелами»), которые в общей сложности охватывают 11 аминокислотных сайтов. Таким образом, $n_{aa}=65+74=139$ и $d_{aa}=74$; следовательно, $p_d=74/139=0,532$ и $K_{aa}=-\ln(1-p_d)=0,760$, причем $\sigma_K=0,091$. Сравнивая таким же образом α -цепи гемоглобина человека и карпа и исключив при

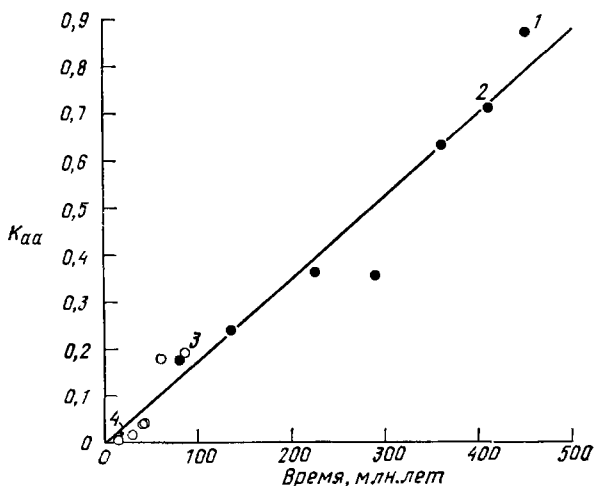


Рис. 4.4. Зависимость между K_{aa} (числом аминокислотных замен) и временем дивергенции (в миллионах лет).

этом 3 пробела, мы получим $d_{aa}=68$ и $n_{aa}=140$. В данном случае $K_{aa}=0,666$ и $\sigma_K=0,082$.

На рис. 4.4 приведен график, на котором черные точки соответствуют расчетным значениям числа аминокислотных замен (K_{aa}), приуроченным к соответствующим временам дивергенции T . Например, точка под номером 1 отвечает среднему значению K_{aa} , полученному при сравнении α -цепи гемоглобина акулы с α -цепями гемоглобинов семи других животных. Аналогичным образом точка 2 — это среднее K_{aa} , полученное при сравнении α -цепи гемоглобина карпа с α -цепями гемоглобинов тритона, курицы и т. д. Из рис. 4.4 видно, что семь черных точек довольно хорошо ложатся на прямую. Наклон прямой равен $0,00179$; следовательно, $k_{aa}=K_{aa}/(2T) \approx 0,9 \cdot 10^{-9}$ на аминокислотный сайт в год. Другими словами, темпы аминокислотных замен в разных ветвях родословного дерева позвоночных были примерно одинаковыми, а репрезентативная величина скорости замещения, равная $0,9 \cdot 10^{-9}$, близка к 10^{-9} . Значительное отклонение наблюдается только при сравнении курицы с млеко-

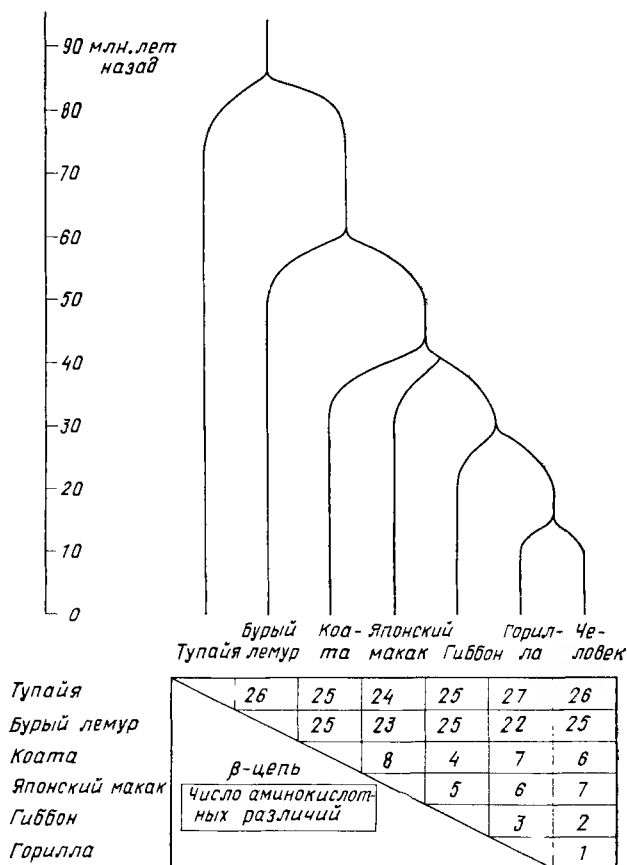


Рис. 4.5. Филогенетическое древо приматов (включая тупайю) и матрица чисел аминокислотных различий, полученных при сравнении β -цепей их гемоглобинов (146 аминокислотных остатков). Используемая при этом расцетке аминокислотная последовательность β -цепи гемоглобина тупайи была установлена Маитой и др. [325].

питающими, что свидетельствует в пользу предположения о более медленной эволюции α -цепей гемоглобина птиц, если, конечно, здесь нет ошибки в определении времени дивергенции.

Кроме того, на графике рис. 4.4 нанесены рассчитанные таким же образом — с использованием данных, приведенных на рис. 4.5, — наблюдаемые значения K_{aa} для β -цепей гемоглобинов приматов (включая тупайю). Эти величины представлены светлыми точками. Точка под номером 3 отражает результат сравнения тупайи с остальными приматами, а точка 4 — резуль-

тат сравнения человека и гориллы. Если судить по приведенным данным, скорость эволюции β -цепей гемоглобина приматов по существу равна скорости эволюции α -цепей гемоглобина позвоночных, т. е. β -цепи гемоглобина приматов эволюционировали со скоростью, примерно равной 10^{-9} на аминокислотный сайт в год.

Единственной кажущейся аномалией является скорость замен для линии, соединяющей человека и гориллу: она много меньше, чем для других эволюционных линий, например в родословной тупайи. В этой связи оценим вероятность того, что число замен, накопившихся за 15 млн. лет (таково время дивергенции человека и гориллы), будет меньше или равно единице в предположении, что скорость замен для β -цепи, состоящей из 146 аминокислотных сайтов, составляет 10^{-9} на сайт в год. Эта вероятность задается двумя первыми членами ряда Пуассона со средней, равной 146 K_{aa} , т. е.

$$P = e^{-X}(1 + X), \quad (4.7)$$

где $X = 146 K_{aa}$. Отметим, что сумма пуассоновских переменных тоже является пуассоновской переменной. Приняв $K_{aa} = 2T k_{aa} = 2 \cdot 1,5 \cdot 10^7 \cdot 10^{-9} = 0,03$, получаем $P = 0,067$ (т. е. $1/15$). Таким образом, у нас нет оснований для окончательного утверждения, что в линии человек — горилла скорость эволюции β -цепи была значимо ниже, чем в других филогенетических линиях позвоночных.

Сходные оценки были получены при сравнении α - и β -цепей гемоглобинов шести млекопитающих — человека, быка, лошади, собаки, кролика и мыши: для α -цепи $k_{aa} = (1,12 \pm 0,09) \cdot 10^{-9}$, а для β -цепи — $(1,28 \pm 0,14) \cdot 10^{-9}$. Здесь принимается, что дивергенция этих животных произошла 80 млн. лет назад. Кроме того, сравнивая миоглобины (состоящие из 153 аминокислот) шести млекопитающих — человека, обыкновенного тюленя, барсука, лошади, быка и кашалота, — мы получаем скорость замен на сайт в год $k_{aa} = (1,01 \pm 0,09) \cdot 10^{-9}$, которая очень близка к приведенным выше оценкам.

Можно предположить, что эволюция глобинов позвоночных происходила с примерно постоянной (одинаковой) скоростью, равной 10^{-9} на сайт в год. Как будет подробно показано в следующем подразделе, такое постоянство наблюдается для скорости, измеряемой числом замен в год; влияние же продолжительности поколения на скорость аминокислотных замен, вообще говоря, невелико.

Примерное постоянство скорости молекулярной эволюции и случайная природа аминокислотных замен обнаруживаются также при сравнении цитохромов с. Цитохром с выполняет важную функцию, участвуя в переносе электронов по дыхательной цепи, и присутствует у большего числа эукариот (т. е. организмов с

оформленным ядром), чем гемоглобин. На рис. 4.6 приведены доли аминокислотных различий между цитохромами с пяти животных, с одной стороны, и цитохромом с пшеницы — с другой. Обратите внимание, что все эти величины, полученные при сравнении различных животных с пшеницей и отражающие степень их дифференциации, почти одинаковы. Однако в данном случае скорость аминокислотных замен в год примерно втрое меньше скорости аминокислотных замен для гемоглобина.

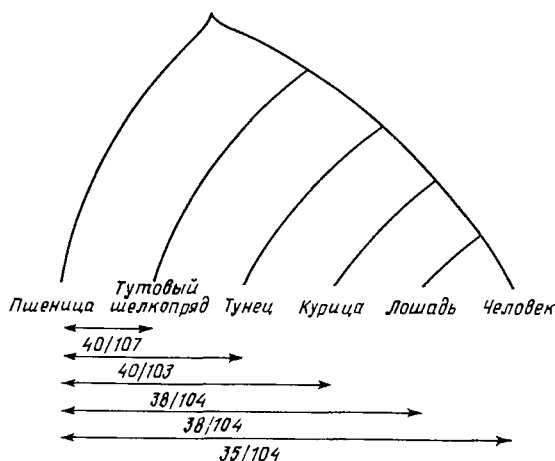


Рис. 4.6. Филогенетическое древо цитохрома с (см. текст). При подсчете числа различий участки, не имеющие соответствия в одной из сопоставляемых последовательностей («пробелы») и возникшие в результате вставок или делеций, не рассматривались.

В табл. 4.1 приведены значения k_{aa} для ряда белков. Самая высокая скорость эволюции отмечается у фибринопептидов ($k_{aa}=8,3 \cdot 10^{-9}$), а самая низкая — у гистона H4 ($k_{aa}=0,01 \cdot 10^{-9}$). Насколько представительны эти величины для всех цистронов (генов) млекопитающих, пока сказать трудно. Возможно, распределение частот этих скоростей не является нормальным, и поэтому в качестве репрезентативной величины больше подходит медиана, а не среднееарифметическое. Оказывается, однако, что медиана скоростей эволюции двадцати одного белка из табл. 6.1 сводки Мак-Лафлина и Дэйхоф [234] равна $1,3 \cdot 10^{-9}$ на аминокислотный сайт в год [235]. Используя табл. 1 из атласа Дэйхоф [77], в которой приведены скорости эволюции шестидесяти белков, мы получаем сходную величину медианной скорости, равную $0,74 \cdot 10^{-9}$ на аминокислотный сайт в год. Эта величина сравнима со средней скоростью для семи белков, полученной Кингом и Джуксом [268] и составляющей $1,6 \cdot 10^{-9}$ на аминокислотный сайт в год. Итак, можно сказать, что стандарт-

Таблица 4.1. Скорости эволюции белков, выраженные в единицах, равных 10^{-9} аминокислотных замен на аминокислотный сайт в год («полингах»), и полученные главным образом по данным о белках млекопитающих¹

Белок	$k_{aa} \cdot 10^{-9}$ в год
Фибринопептиды	8,3
Панкреатическая рибонуклеаза	2,1
Лизоцим	2,0
α -Цепь гемоглобина	1,2
Миоглобин	0,89
Инсулин	0,44
Цитохром <i>c</i>	0,3
Гистон H4	0,01

¹) Все величины взяты из работы Дэйхоф [77]. Исключение составляют: скорость эволюции лизоцима, полученная по данным Уилсона и др. [584] с использованием соотношения k_{aa} (в полингах) = $5/EB\dot{E}$ (в миллионах лет); скорость эволюции фибринопептидов, рассчитанная Отой и Кимурой [453], которые изучали быстро эволюционировавшие участки фибринопептидов А и В; скорость эволюции цитохрома *c*, полученная Мак-Лафлином и Дэйхоф [324].

ная скорость молекулярной эволюции (скорость аминокислотных замен) примерно равна 10^{-9} на сайт в год и что темпы эволюции глобинов позвоночных вполне репрезентативны. В качестве единицы скорости эволюции белков я предложил [231] использовать величину 10^{-9} на аминокислотный сайт в год, назвав ее *полингом*. Скорость эволюции гемоглобина, выраженная в этих единицах, очень близка к одному полингу, а скорость эволюции цитохрома *c* составляет около 30 сантиполингов. К сожалению, этот термин не стал популярным у исследователей, изучающих молекулярную эволюцию, точно так же как в работах по фенотипической эволюции не привился термин *дарвин*. По сравнению с более популярными единицами, такими, как «число точковых мутаций (accepted point mutation) на 100 аминокислотных остатков, накопившееся за 100 млн. лет» (см. с. 3 атласа Дэйхоф [77]), «единичный временной интервал эволюции» (Unit Evolutionary Period), или сокращенно ЕВЭ (определяемый как «время в миллионах лет, необходимое для возникновения различия в 1% между аминокислотными последовательностями двух дивергировавших линий» [80]), полинги представляют гораздо более простую и прямую меру скоростей молекулярной эволюции, и поэтому я надеюсь, что в будущем они найдут более широкое применение. Может быть, вместо полинга следовало бы ввести какой-нибудь «безличный» термин и называть эту единицу, скажем, «единицей молекулярной эволюции» (Molecular Evolutionary Unit).

Зависимость, задаваемая уравнением (4.4) и связывающая наблюдаемое число аминокислотных различий между двумя гомологичными последовательностями с расчетным числом аминокислотных замен, является достаточно хорошим приближением, если доля различающихся аминокислот не слишком велика (скажем, не превышает 50%). Более точное соотношение между этими двумя величинами, представленное в виде переводной таблицы, получено Дэйхоф (см. табл. 23 в работе [77]). При составлении этой таблицы использовалась «матрица вероятностей мутирования». Это позволило учесть то обстоятельство, что для некоторых аминокислот «взаимозамещаемость» гораздо выше, чем при случайном осуществлении замен. По сравнению с зависимостью, полученной Дэйхоф, уравнение (4.4) приводит к заниженным оценкам числа аминокислотных замен, особенно в случае больших эволюционных расстояний.

Приведенная ниже эмпирическая формула, полученная путем простого преобразования соотношения (4.4), дает зависимость между K_{aa} и p_d , весьма близкую к той, которая следует из таблицы Дэйхоф:

$$K_{aa} = -\ln \left(1 - p_d - \frac{1}{5} p_d^2 \right). \quad (4.8)$$

Как видно из табл. 4.2, эта формула справедлива при значениях p_d , не превышающих 0,7. Синонимичные замены нуклеотидов в этих расчетах не учитывались. Если бы частоты нуклеотидных замен во всех трех положениях кодона были одинаковы, синонимичные замены нуклеотидов составляли бы прибли-

Таблица 4.2. Зависимость между долей аминокислотных сайтов, по которым наблюдаются различия (p_d), и расчетным числом аминокислотных замен (K_{aa})

p_d	K_{aa}		
	по Дэйхоф ¹⁾	формула (4.4)	формула (4.8)
0,01	0,01	0,01	0,01
0,05	0,05	0,05	0,05
0,10	0,11	0,11	0,11
0,20	0,23	0,22	0,23
0,30	0,38	0,36	0,38
0,40	0,56	0,51	0,57
0,50	0,80	0,69	0,80
0,60	1,12	0,92	1,11
0,70	1,59	1,20	1,60
0,80	2,46	1,61	2,63

¹⁾ Из табл. 23 атласа Дэйхоф [77].

зительно 25% [230], однако в действительности такие замены, как будет показано ниже, происходят гораздо чаще.

Следует добавить, что существует несколько методов, подобных методу минимального филетического расстояния Фитча [114], позволяющих оценить число нуклеотидных замен из данных по аминокислотным последовательностям. Однако, как было показано Неем и Чакраборти [407], оценки, полученные разными методами, достаточно близки, и поэтому на практике для построения эволюционных древ или датировки нуклеотидных замен могут использоваться любые из них.

Постоянство темпов молекулярной эволюции

По сравнению с эволюцией на фенотипическом уровне молекулярная эволюция имеет одну замечательную особенность, состоящую в том, что ее скорость примерно постоянна (т. е. одинакова) в различных родословных данного белка. Мы уже отмечали этот замечательный феномен (см. предыдущий подраздел), однако ввиду того, что он имеет самое непосредственное отношение к теории нейтральности, я рассмотрю его более подробно.

В табл. 4.3 приведены наблюдаемые числа аминокислотных различий между α -цепями гемоглобинов шести млекопитающих (человека, мыши, кролика, собаки, лошади и быка). Эти млекопитающие дивергировали, по-видимому, в конце мезозоя, около 80 млн. лет назад, и с тех пор в их генах, кодирующих α -цепь гемоглобина, независимо накопилось какое-то количество мутаций. Посмотрим, одинаковы ли скорости аминокислотных замен для этих шести родословных и укладывается ли вариабельность наблюдаемых чисел замен, приведенных в табл. 4.3, в рамки обычных статистических флуктуаций.

Пусть $x_1, x_2, \dots, x_n$ — число аминокислотных замен в n филогенетических линиях, накопившееся после их дивергенции от общего предка А (см. рис. 4.7, где рассмотрен случай $n=6$).

Таблица 4.3. Наблюдаемое число аминокислотных различий между α -цепями гемоглобинов шести млекопитающих. Отношение наблюдаемых и ожидаемых дисперсий равно $R=1,26$

Человек	Мышь	Кролик	Собака	Лошадь	Бык
Человек	18	25	23	18	17
Мышь		27	25	24	19
Кролик			28	25	25
Собака				27	28
Лошадь					18
Бык					

Обозначим среднее и дисперсию x через $\bar{x}$ и V_x . Пусть $\bar{D}$ и V_D — среднее и дисперсия числа аминокислотных различий, выявляемых при сравнении n видов, т. е. при проведении $n(n-1)/2$ сравнений. Можно показать, что $\bar{x} = \bar{D}/2$ и $V_x = (n+1)V_D/[2(n-1)]$. Среднее и дисперсия для 15 наблюдаемых величин [после введения пуассоновской поправки с использованием уравнения (4.4) и с учетом того, что $n_{aa}=141$], приведенных в табл. 4.3, оказываются равными $\bar{D}=25,34$ и $V_D=22,81$. Следовательно, $\bar{x}=12,67$ и $V_x=15,97$.

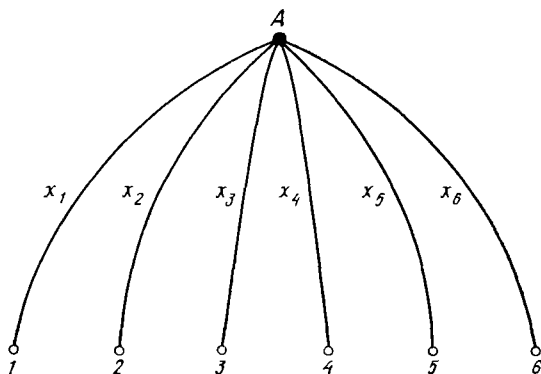
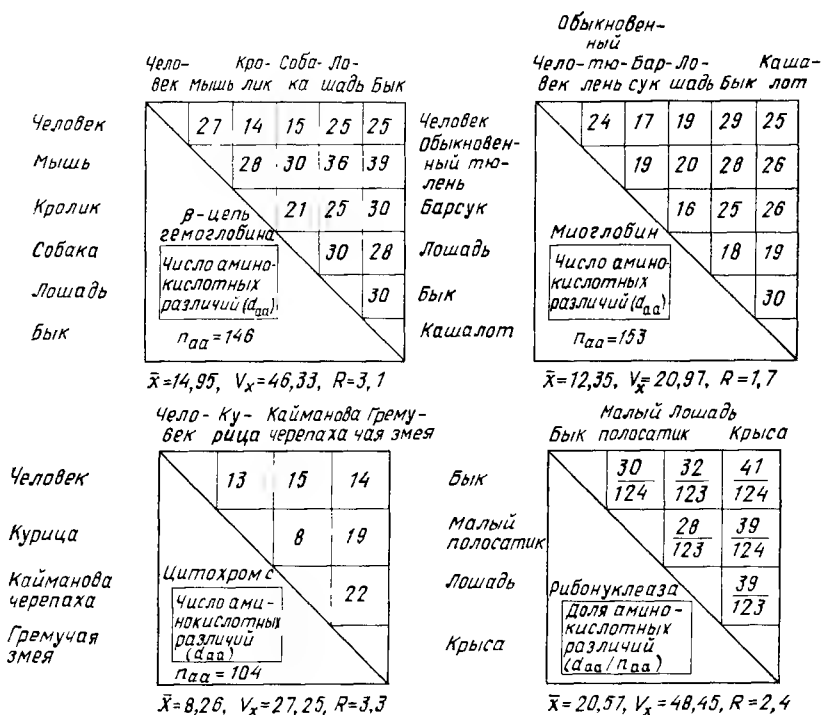


Рис. 4.7. Дивергенция шести млекопитающих (см. табл. 4.3) от их общего предка А. x_1 , x_2 и т. д. — числа аминокислотных замен в различных филогенетических линиях. Для данных из табл. 4.3 среднее и дисперсия равны $\bar{x}=12,67$ и $V_x=15,97$.

Чтобы установить, превышает ли вариабельность числа аминокислотных замен по разным родословным вариабельность, ожидаемую из чисто стохастических соображений, достаточно найти величину $R = V_x/\sigma_x^2$, где σ_x^2 — теоретически ожидаемая дисперсия. R , как видно из формулы, представляет собой отношение наблюдаемой и ожидаемой дисперсий. Так как x_i следуют распределению Пуассона и так как при этом распределении дисперсия равна среднему, мы можем принять $\sigma_x^2 = \bar{x}$, и тогда для рассматриваемого случая получим $R=1,26$. Статистика $(n-1)R$ должна следовать распределению χ^2 с $n-1$ степенями свободы. В нашем случае $n=6$ и $(n-1)R=6,3$. Из таблицы распределения χ^2 следует, что отклонение R от единицы статистически недостоверно, и, таким образом, мы не можем считать, что скорости аминокислотных замен в этих шести филогенетических линиях различаются. Следует отметить, что если точки дивергенции шести рассматриваемых родословных не совпадают, R , как правило, дает завышенное по сравнению с истинным значение отношения дисперсий, и поэтому описанный

Таблица 4.4. Иллюстрация теста на постоянство скоростей молекулярной эволюции в разных родословных применительно к β -цепи гемоглобина, миоглобину, цитохрому c и панкреатической рибонуклеазе (описание этого теста на примере α -цепи гемоглобина дается в тексте). Данные взяты из атласа Дэйхоф [77].

Длина полипептидной цепи гемоглобина $n_{aa}=146$, миоглобина — 153 и цитохрома c — 104. В случае рибонуклеазы $n_{aa}=123$ или 124 в зависимости от того, какие последовательности сравниваются, так как при их сопоставлении выявляются «пробелы» (т. е. вставки или делеции), исключаемые при расчете долей аминокислотных различий.



тест на постоянство темпов молекулярной эволюции является консервативным.

Тот же тест можно применить и к β -цепи гемоглобина, цитохрому c и панкреатической рибонуклеазе (см. табл. 4.4). Считается, что животные, сравниваемые по каждой из этих молекул, дивергировали от общего предка примерно в одно и то же время, и поэтому продолжительность их раздельной эволюции была одинаковой. Оказывается, что для β -цепи гемоглобина $R=3,1$, для миоглобина — 1,7, для цитохрома c — 3,3, а для ри-

Таблица 4.5. Число аминокислотных различий, выявленных при сравнении α - и β -цепей разных гемоглобинов

Тип аминокислотного различия ¹⁾	Число аминокислотных различий			
	α -цепь человека — β -цепь человека	α -цепь карпа — β -цепь человека	α -цепь утконоса — β -цепь утконоса	α -цепь акулы — β -цепь акулы
0	62	61	62	50
1	55	49	49	56
2	21	29	28	32
3	0	0	0	1
Пробел	9	10	9	11
Сумма	147	149	148	150

¹⁾ Аминокислотные различия разных типов (обозначенных через 0, 1, 2 и 3), выявленные в этих четырех сериях сравнений α - и β -цепей гемоглобинов, обусловлены, как следует из таблицы генетического кода, как минимум, 0, 1, 2 и 3 нуклеотидными заменами соответственно (см. табл. 4.6). Для каждого случая приводится также число пробелов (выраженное через эквивалентное число аминокислот).

бонуклеазы $R=2,4$. Отклонение от единицы значимо для β -цепей гемоглобина (на уровне примерно 1%) и цитохромов c (на уровне 3%), но незначимо в двух других случаях. Приведенные результаты свидетельствуют о том, что постоянство — если не строгое, то приблизительное — скорости эволюции любой данной макромолекулы в разных родословных является скорее правилом, чем исключением. Следует отметить, что разные макромолекулы могут эволюционировать с разными скоростями. Так, например, скорость эволюции панкреатической рибонуклеазы у млекопитающих ($k_{aa}=2,07 \cdot 10^{-9}$) оказывается примерно в восемь раз выше, чем цитохрома c у позвоночных ($k_{aa}=0,26 \cdot 10^{-9}$). Среднее значение R для пяти упомянутых выше белков (включая α -цепь гемоглобина) равно 2,6. Это согласуется с результатами нашего давнего исследования [454], в котором мы показали, что наблюдаемые дисперсии скоростей эволюции гемоглобинов и цитохрома c в филогенетических линиях млекопитающих примерно в 1,5—2,5 раза больше дисперсии, теоретически ожидаемой для ситуации, когда вариабельность этих скоростей определяется чисто случайными процессами.

Наблюдается ли такое же примерное постоянство скоростей молекулярной эволюции при сравнении более далеких таксонов? Рассмотрим дивергенцию α - и β -цепей гемоглобина. Эти две полипептидные цепи появились в результате дупликации, которая произошла в отдаленном прошлом до появления костистых рыб. Во втором и третьем столбцах табл. 4.5 сравниваются дивергенция α - и β -цепей гемоглобина человека и дивергенция α -цепи гемоглобина карпа и β -цепи гемоглобина человека. Эта

Таблица 4.6. Минимальное число замен нуклеотидов («мутационных шагов»), необходимых для перехода от одной аминокислоты к другой. (Например, для перехода от треонина к валину требуются по крайней мере две нуклеотидные замены: A→G в первом положении кодона и G→U во втором.)

	Ala	Cys	Asp	Glu	Phe	Gly	His	Ile	Lys	Leu	Met	Asn	Pro	Gln	Arg	Ser	Thr	Val	Trp	Tyr
	A	C	D	E	F	G	H	I	K	L	M	N	P	Q	R	S	T	V	W	Y
Ala A	0	2	1	1	2	1	2	2	2	2	2	1	2	2	1	1	1	2	2	
Cys C	2	0	2	3	1	1	2	2	3	2	3	2	2	3	1	1	2	2	1	1
Asp D	1	2	0	1	2	1	1	2	2	2	3	1	2	2	2	2	2	1	3	1
Glu E	1	3	1	0	3	1	2	2	1	2	2	2	2	1	2	2	2	1	2	2
Phe F	2	1	2	3	0	2	2	1	3	1	2	2	2	3	2	1	2	1	2	1
Gly G	1	1	1	1	2	0	2	2	2	2	2	2	2	2	1	1	2	1	1	2
His H	2	2	1	2	2	2	0	2	2	1	3	1	1	1	1	2	2	2	3	1
Ile I	2	2	2	2	1	2	2	0	1	1	1	2	1	1	1	1	1	1	3	2
Lys K	2	3	2	1	3	2	2	1	0	2	1	1	2	1	1	2	1	2	2	2
Leu L	2	2	2	2	1	2	1	1	2	0	1	2	1	1	1	1	2	1	1	2
Met M	2	3	3	2	2	2	3	1	1	1	0	2	2	2	1	2	1	1	2	3
Asn N	2	2	1	2	2	2	1	1	1	2	2	0	2	2	2	1	1	2	3	1
Pro P	1	2	2	2	2	2	1	2	2	1	2	2	0	1	1	1	1	2	2	2
Gln Q	2	3	2	1	3	2	1	2	1	1	2	2	1	0	1	2	2	2	2	2
Arg R	2	1	2	2	2	1	1	1	1	1	1	2	1	1	0	1	1	2	1	2
Ser S	1	1	2	2	1	1	2	1	2	1	2	1	1	2	1	0	1	2	1	1
Thr T	1	2	2	2	2	2	2	1	1	2	1	1	1	2	1	1	0	2	2	2
Val V	1	2	1	1	1	1	2	1	2	1	1	2	2	2	2	2	2	0	2	2
Trp W	2	1	3	2	2	1	3	3	2	1	2	3	2	2	1	1	2	2	0	2
Tyr Y	2	1	1	2	1	2	1	2	2	3	1	2	2	2	1	2	2	2	0	2

таблица составлена по данным атласа Дэйхоф ([77]; Alignment 32) с помощью табл. 4.6, в которой приведены минимальные числа нуклеотидных замен, необходимых для перехода от одной аминокислоты к другой, для 20·20 комбинаций аминокислот.

Так как α -цепи гемоглобинов человека и карпа различаются примерно по 50% аминокислотных сайтов (см. таблицу в нижней части рис. 4.2), приведенные данные свидетельствуют о том, что в этих двух родословных протекала независимая и симметричная дивергенция двух генов, кодирующих α - и β -цепи, после того как эти гены появились в результате дупликации. Примечательно, что мутационные замены происходили с практически одинаковой средней скоростью в двух разделившихся филогенетических линиях, независимо эволюционировавших на протяжении почти полумиллиарда лет. Для сравниваемых последовательностей скорость аминокислотных замен оказывается рав-

ной примерно $0,81 \cdot 10^{-9}$, если пользоваться формулой (4.4) (с пуассоновской поправкой), и $0,96 \cdot 10^{-9}$ согласно формуле (4.8) (с поправкой Дэйхоф), если принять, что дупликация α - и β -цепей произошла около $5 \cdot 10^8$ лет назад. Так как сопоставляемые последовательности очень сильно различаются, для этого расчета, вероятно, больше подходит поправка Дэйхоф. Тем не менее обе оценки темпов аминокислотных замен довольно близки к $0,9 \cdot 10^{-9}$ — величине, полученной в предыдущем подразделе при сравнении последовательностей α -цепей гемоглобинов позвоночных. К сходным результатам приводит сравнение α - и β -цепей гемоглобинов других животных, например α -2- и β -цепей гемоглобина курицы или α - и β -цепей гемоглобина утконоса.

Эти факты свидетельствуют о том, что скорость молекулярной эволюции, выраженная через число мутационных замен, постоянна в расчете на год и, как это ни удивительно, по-видимому, не зависит от таких факторов, как продолжительность поколения, условия жизни и размер популяции. Как отмечалось в предыдущей главе, для селективно нейтральных мутантов скорость эволюции, выраженная через число мутационных замен, равна скорости мутирования на гамету и не связана с численностью популяции [см. формулу (3.20)]. Таким образом, самое простое объяснение наблюдаемого постоянства темпов молекулярной эволюции белков основано на предположении, что большинство аминокислотных замен — результат случайной фиксации селективно нейтральных мутаций и что скорость появления нейтральных мутаций на гамету в расчете на год постоянна. В данной связи я в свое время заметил [231], что если в генах гемоглобинов и других белков живых ископаемых накопилось столько же замен оснований (а в самих белках — аминокислотных замен), сколько в соответствующих генах (белках) более быстро эволюционировавших видов, то это может рассматриваться как аргумент в пользу моей теории нейтральности.

Недавно были установлены аминокислотные последовательности β - и α -цепей основного гемоглобина акулы Порт Джексона (*Heterodontus portusjacksoni*) [111]. Согласно Ромеру [493], эта акула является уцелевшим реликтовым видом и принадлежит к группе предковых акул, многочисленные представители которой обитали в позднем палеозое и особенно в каменноугольном периоде (270—350 млн. лет назад). Таким образом, эту акулу можно с полным основанием назвать живым ископаемым. В крайнем правом столбце табл. 4.5 приведены результаты сравнения α - и β -цепей гемоглобина акулы Порт Джексона. Если сопоставить эти результаты с соответствующими данными для человека, то станет ясно, что гены, кодирующие α - и β -цепи гемоглобина акулы и возникшие в результате дупликации, ди-

вергировали вследствие накопления случайных мутаций приблизительно в такой же (или несколько меньшей) степени, что и соответствующие два гена человека.

Другой пример живого ископаемого представляет опоссум. Чтобы проверить, правильно ли мое предположение относительно скорости молекулярной эволюции живых ископаемых, Стензел [531] секвенировал α -цепь гемоглобина этого животного. Из его данных (см. табл. 2 в работе [531]) следует, что α -цепь опоссума со времени его дивергенции от плацентарных млекопитающих, происшедшей около 135 млн. лет назад, эволюционировала со скоростью $k_{aa}=1,7 \cdot 10^{-9}$ в год. Таким образом, как отметил Стензел [531], «полученное значение k_{aa} превышает значение, ожидаемое на основе теории нейтральности». Я хотел бы заметить, что первичная структура α -цепи гемоглобина опоссума, секвенированной Стензелом, имеет одну совершенно необычную особенность. Известно [200], что у позвоночных в положениях 58 и 87 α -цепи гемоглобина находится инвариантный остаток His (связанный с гемом), однако в α -цепи гемоглобина опоссума положение 58 занято не His, а Gln. Такой гемоглобин можно назвать патологическим. Вероятно, в ходе эволюции опоссума по неизвестным причинам (например, в результате дупликации гена) были сняты некоторые селективные ограничения, налагаемые на первичную структуру α -цепи гемоглобина, за которым последовало необычайно быстрое накопление мутаций. Если это верно, то причина высокой скорости замен аминокислот в α -цепи гемоглобина опоссума сходна с той, которая вызвала быстрое накопление аминокислотных замен в инсулине морской свинки, связанное с потерей обычно инвариантного гистидина B10 (эволюция этого белка будет рассмотрена в следующей главе). Что же касается последовательности миоглобина опоссума ([76]; Alignment 33), то в ней нет ничего необычного. Скорость эволюции этого белка равна $k_{aa}=0,6 \cdot 10^{-9}$ в год, что немного меньше обычной для миоглобина величины $0,9 \cdot 10^{-9}$ в год.

Проблема постоянства эволюционной скорости принадлежит к числу наиболее дискуссионных проблем молекулярной эволюции; особенно острую полемику вызвала ее интерпретация с позиций теории нейтральности. Тезис о постоянстве скорости молекулярной эволюции и его нейтралистская интерпретация критиковались многими авторами. Наиболее видный из них — Левонтин — в своей книге «Генетические основы эволюции» ([302], см. с. 229) высказывает мнение, что «сторонники концепции «постоянства» просто-напросто перепутали среднюю и постоянную величины». Он указывает на то, что говорить о постоянстве скорости молекулярной эволюции — это то же самое, что утверждать, будто температура в Чикаго абсолютно неизменна, на том основании, что, измеряя ее в течение послед-

них пятидесяти лет прошлого века и первых пятидесяти — нынешнего, получили одну и ту же величину. Еще раньше Стеббинс и Левонтин [529] писали, что «так называемое постоянство скоростей эволюции биологических макромолекул в течение миллионов лет есть просто следствие закона больших чисел». Подобная критика стала довольно популярной (см., например, [171, 524]) и даже приводилась в учебнике эволюционной биологии Гранта [142], который отверг мою идею о постоянстве темпов молекулярной эволюции как основанную на «элементарной ошибке».

Здесь я хотел бы заметить, что критика Левонтина основана только на неправильном понимании действительной природы молекулярной эволюции. Обратимся к примеру с α -цепью гемоглобина. Как мы уже знаем, при эволюции α -цепи гемоглобина, протекающей в какой-либо филогенетической линии, одна аминокислотная замена происходит в среднем примерно каждые 7 млн. лет. Ясно, что для получения адекватной оценки скорости эволюции этого белка необходимо учесть число замен, происшедших на протяжении по крайней мере нескольких миллионов лет, а лучше — в течение значительно более длительных временных интервалов. Рассматриваемая ситуация имеет определенное сходство со статистикой несчастных случаев. В такой ситуации действует скорее «закон малых чисел» (закон Пуассона), а не больших. Действительно важно здесь, однако, то, что в соответствующих сравнениях «участвуют» две или большее число различных родословных, подобных изображенным на рис. 4.2, 4.4 и 4.5 (см. также табл. 4.3 и 4.4), и для проведения этих сравнений мы должны оценить «истинные» (intrinsic) скорости эволюции для разных филогенетических линий. Как не может стать одинаковой при простом усреднении за длительный период или для большого числа особей смертность для человека и какого-нибудь насекомого, так и две средние величины не станут конвергировать, если за ними не стоят одинаковые «внутренние» (intrinsic) параметры. По моему убеждению, «истинные» скорости молекулярной эволюции определяются не условиями обитания, а структурой и функциями соответствующих молекул.

При обсуждении обозначенной проблемы следует проводить четкое различие между наблюдаемой скоростью эволюции и «внутренним» параметром, лежащим в ее основе. Рассмотрим, как в ходе эволюции некоего вида млекопитающих осуществляется процесс мутационных замен, приводящих к заменам аминокислот, в гене, кодирующем α -цепь гемоглобина. Если каждое поколение данного вида представлено полумиллионом особей ($N=0,5 \cdot 10^6$) и если длительность поколения равна двум годам, то общее число мутаций, которое произойдет в данном локусе за 7 млн. лет при скорости мутирования в расчете на поколение

и на гамету, равной 10^{-6} , составит 3,5 млн. Примечательно, что в геноме вида фиксируется лишь одна из этих 3,5 млн. мутаций. Наряду с прямыми мутациями осуществляется, конечно, и некоторое количество обратных мутаций, а кроме того, не все появляющиеся мутации селективно нейтральны, однако для нашего анализа это несущественно.

Учитывая все сказанное выше и принимая тезис о стохастичности процесса замещения аминокислот, нетрудно понять, что «истинная» скорость замен и наблюдаемая частота замен — разные вещи. По мере «продвижения» в глубь длительной эволюционной истории млекопитающих число наблюдаемых замен, происшедших за определенный отрезок времени, скажем за 10 млн. лет, должно сильно флуктуировать, однако отсюда не следует однозначный вывод о флуктуации самой «истинной» молекулярной эволюции. За какой-то один временной интервал может не произойти ни одной замены, а за другой интервал такой же продолжительности могут случайным образом произойти две замены или даже более. Эту ситуацию можно сравнить с серией испытаний, в каждом из которых одновременно бросают 20 монет. «Герб» на *всех* 20 монетах может выпасть примерно в одном из миллиона испытаний, хотя «истинная» вероятность такого события всегда одинакова.

Факт постоянства (или примерного постоянства) скорости молекулярной эволюции во времени (измеряемом в годах) нашел отражение в концепции «молекулярных эволюционных часов», сформулированной Цукеркандлом и Полингом [622]. С тех пор термин «молекулярные часы» приобрел большую популярность у исследователей, работающих над проблемами молекулярной эволюции.

Эта концепция была с энтузиазмом воспринята и поддержана Уилсоном и др. [584], которые писали, что «открытие эволюционных часов оказалось самым значительным результатом, полученным при изучении молекулярной эволюции». Однако, по их мнению, дискуссия нейтралистов и селекционистов не имеет первостепенного значения для выяснения механизма работы эволюционных часов. Касаясь этой дискуссии, они выражают несогласие с мнением исследователей, считающих решение дилеммы нейтрализма и селекционизма главной задачей, стоящей перед специалистами в области молекулярной эволюции. Уилсона и др. интересуется не механизм эволюционных часов, а сами часы и их использование для расстановки «верстовых столбов» эволюции. Кстати, Фитч и Лэнгли [116] писали, что «эти часы время от времени идут неритмично, однако при подсчете большого числа «тиканий» их точность оказывается удовлетворительной». Утверждая, что в случае больших выборок «две средние величины могут быть одинаково репрезентативными», они, как мне кажется, повторяют ошибку Левонтина.

Тезис о постоянстве скорости эволюции, выраженной через число аминокислотных замен, является научной гипотезой, и согласие или несогласие с ним должно подкрепляться не риторическими рассуждениями, а фактами и их статистическим анализом. Именно по этой причине мы с Отой [454] попытались проанализировать дисперсии скоростей эволюции в разных филогенетических линиях. Мы установили, что наблюдаемая дисперсия несколько выше величины, ожидаемой для случая, когда вариабельность этой скорости обусловлена чисто случайными процессами. Более детальный статистический анализ был проведен Лэнгли и Фитчем [291]. Используя метод максимального правдоподобия и приняв предположение о «минимальной эволюции», они проанализировали вариабельность эволюционной скорости для набора белков, включающего α - и β -цепи гемоглобина, цитохром *c* и фибринопептид А, для разных ветвей филогенетического древа. Оказалось, что вариабельность скорости эволюции для разных родословных значительно выше величины, ожидаемой из чисто вероятностных соображений, причем значение χ^2 примерно вдвое больше числа ее степеней свободы. Так как ожидаемое значение χ^2 равно числу степеней свободы, то это означает, что вариабельность эволюционной скорости, выраженной через число мутационных замен, по которым различаются белки представителей разных филогенетических линий, примерно вдвое больше ожидаемой величины. Этот результат Лэнгли и Фитч расценили как свидетельство непостоянства скорости молекулярной эволюции и сделали вывод, что «значительная часть наблюдаемых замен зафиксирована благодаря отбору и, по-видимому, связана с изменением окружающей среды. Предположение о постоянстве относительных скоростей селективных замен во времени или в разных филогенетических линиях лишено оснований». Другими словами, Лэнгли и Фитч не считают нейтральные мутации главной причиной появления замен. Однако в той же работе они показали, что если отложить расчетные числа мутационных замен в зависимости от соответствующих геологических дат, то получится прямая (аналогичная прямой на рис. 4.4), а это свидетельствует о постоянстве темпов молекулярной эволюции. Не ложатся на прямую только три точки, отвечающие приматам. По моему мнению, считать локальные флуктуации свидетельством против теории нейтральности, не пытаясь выяснить причины столь замечательного постоянства скорости молекулярной эволюции в целом, совершенно неоправданно. Это классический пример ситуации, когда «за деревьями не видят леса». Замечательное постоянство скоростей молекулярной эволюции (в расчете на год) было обнаружено также Уилсоном и его группой в ходе иммунологического изучения альбуминов разных позвоночных [584].

Тезис о непостоянстве скорости молекулярной эволюции раз-

вивали Гудмен и др. [138, 139] в своих работах, посвященных эволюции глобина. Согласно этим авторам, на ранних этапах эволюции глобина, особенно вскоре после дупликации, приведшей к образованию миоглобина, а также α - и β -цепей гемоглобина, темп мутационных замен был очень высок, а затем на протяжении последних 300 млн. лет, в ходе эволюции от амниотического предка до современных организмов он заметно снизился. По Гудмену и др. [139], за период 500—400 млн. лет назад гены глобинов, которые произошли от общего для всех позвоночных предкового гена, последовательно превращавшегося в гены-предшественники гемоглобина и миоглобина, α - и β -цепей гемоглобина и α -цепи гемоглобина костистых рыб и тетрапод, эволюционировали со скоростью 109 NR% (где NR% — число нуклеотидных замен на 100 кодонов за 10^8 лет). Приведенная величина, если следовать терминологии, принятой в нашей книге, соответствует $k_{aa}=10,9 \cdot 10^{-9}$, что примерно в 10 раз больше полученной нами оценки, равной 10^{-9} на аминокислотный сайт в год. С другой стороны, за последние 300 млн. лет эта скорость, по оценкам Гудмена и др., была равна только 15 NR% (т. е. $k_{aa}=1,5 \cdot 10^{-9}$), что не очень отличается от $k_{aa}=1,12 \cdot 10^{-9}$ — оценки, полученной нами при сравнении α -цепей глобинов ряда млекопитающих. Тезис Гудмена и др. [139] о непостоянстве скорости молекулярной эволюции наряду с предложенной ими правдоподобной концепцией, согласно которой появление у глобинов новых функций сопровождалось быстрыми мутационными заменами в дублированных глобиновых генах, вызванными *положительным* дарвиновским отбором, часто приводится в качестве некоего аргумента против теории нейтральности (см., например, [561]).

Если тщательно проанализировать данные, использованные Гудменом и др. [138, 139], то обнаружится, что полученная ими высокая оценка скорости эволюции глобинов (109 NR%) по существу следует из их предположения, согласно которому дупликация гена, приведшая, с одной стороны, к появлению глобина миоглобины, а с другой — предкового глобина, предшественника миоглобина, а также α - и β -глобинов, произошла 500 млн. лет назад (обозначим эту точку филогенетического дерева через *a*; см. рис. 1 в работе [139]). Они также предполагают, что за этим последовала дивергенция миоглобина и (β — α)-гемоглобина, происшедшая примерно 30 млн. лет спустя (обозначим соответствующую точку филогенетического дерева через *b*). Кроме того, Гудмен и др. постулируют, что дупликация, в результате которой появились гены α - и β -цепей гемоглобина, произошла около 425 млн. лет назад (точка *c* на филогенетическом древе) и что предковые гены α -цепи гемоглобинов карпа и курицы дивергировали 400 млн. лет назад (точка *d*). Между точками *a* и *b* они поместили 50 мутаций, между *b* и *c* — 58, а меж-

ду *c* и *d* — 51. Иными словами, согласно их предположениям, за 100 млн. лет (разделяющие точки *a* и *b*) произошли две генные дубликации и в общей сложности 159 мутаций (приходящихся примерно на 145 аминокислотных сайтов).

Принятая ими датировка дубликаций глобиновых генов порождает ряд проблем, самая серьезная из которых связана с их предположением, что дубликация, приведшая к дивергенции миоглобина и гемоглобина, произошла около 470 млн. лет назад (в точке *b*) после отделения предка челюстных рыб от предка миноги. Другими словами, они полагают, что эволюционная история миоглобина короче эволюционной истории бесчелюстных.

Гораздо более вероятно, что дубликация гена, приведшая к появлению глобина миноги (*Petromyzon marinus*), аминокислотная последовательность которого использовалась в их работе, и белок, явившийся общим предком миоглобина и α - и β -цепей глобинов, возникли намного раньше, чем постулируют эти авторы. Другими словами, дублицированные глобиновые гены, по-видимому, существовали уже у предковых агнат, появившихся 500 млн. лет назад (точка *a*). Согласно Хендриксону и Лаву [178], глобин миноги, о котором шла речь выше, представляет собой лишь один из шести гемоглобинов, присутствующих в крови *P. marinus*, а именно гемоглобин V, преобладающий у этих круглоротых. Поэтому можно полагать, что дубликации генов осуществляются в ходе эволюции довольно часто. Но еще более важным для нас представляется недавнее исследование Ромеро-Эрреры и др. [494], обнаруживших настоящий миоглобин в красной мускулатуре сердца миноги. Хотя его аминокислотная последовательность еще не установлена, этот результат полностью опровергает предположение Гудмена и др. о том, что миоглобин появился только около 470 млн. лет назад после отделения челюстноротых от бесчелюстных и что поэтому миоглобин моложе HbV миноги. Гораздо больше осмотрительности проявили Хант и др. [188], расположившие ветвь глобина миноги на противоположной стороне филогенетического дерева по отношению к ветвям других глобинов позвоночных и выдвинувшие тезис о большой древности генной дубликации, приведшей к появлению миоглобина и глобина миноги (на рисунке, приведенном в работе [188], это событие относят примерно к одному миллиарду лет назад). Кроме того, они относят точку дивергенции α - и β -цепей приблизительно к 500 млн. лет назад; эта датировка намного правдоподобнее датировки Гудмена и др., которые считают, что дивергенция α - и β -цепей произошла около 425 млн. лет назад.

Гудмен и др. [137—139] критиковали тезис о постоянстве скорости молекулярной эволюции также на том основании, что использованная мной процедура оценки не учитывает множест-

венных мутаций; они пишут, что их метод расчета якобы намного лучше моего. Давайте рассмотрим именно эту сторону указанных работ. Метод Гудмена состоит, во-первых, в оценке числа мутационных замен с использованием принципа «наибольшей экономии» (principle of „maximum parsimony”), а во-вторых, в коррекции этой оценки — порой очень значительной — с помощью введенной ими процедуры дополнения („augmentation procedure”). Возьмем, например, эволюционные изменения α -цепей гемоглобина в двух родословных — одной, ведущей к карпу, и другой, ведущей к человеку, — начинающихся от их общего предка, который существовал около 400 млн. лет назад. Согласно Гудмену и др. [138], число мутаций, накопленных в родословной карпа, равно 64, а число мутаций, зафиксированных в родословной человека, — 56; обе эти оценки были получены методом наибольшей экономии и представлены в виде «А-решений» („A-solution” values) (см. рис. 2 в работе [138]). Эти две величины, конечно, не свидетельствуют о каком-либо существенном различии между «истинными» скоростями мутационных замен в двух сравниваемых родословных. Далее, они скорректировали полученные оценки, «дополнив» величину, полученную для линии, ведущей к карпу, 48 мутациями, а для линии, ведущей к человеку, — 33 мутациями. Таким образом, окончательные оценки числа мутационных замен, накопленных в этих двух родословных, оказываются равными соответственно 112 и 89 (см. рис. 6 в работе [138]). Хотя статистические ошибки этих оценок неизвестны, есть основания предполагать, что они очень велики, и поэтому различие между двумя приведенными величинами не может считаться достоверным. Это вполне очевидно из того факта, что в другой работе, опубликованной в следующем году [139], они поместили в тех же двух родословных уже не 112 и 89, а соответственно 109 и 105 мутаций (см. рис. 1 в работе [139]). Да и сама процедура дополнения чревата, как отмечали Татено и Ней [545], а также Фитч [115], серьезными статистическими проблемами. К тому же использованный ими принцип наибольшей экономии, по-видимому, ошибочен в самой своей основе, так как у нас нет гарантии, что эволюция протекает кратчайшим путем.

По мнению Гудмена и др. [138], метод наибольшей экономии часто позволяет выявлять различия между разными кодо-нами одной и той же аминокислоты и благодаря столь высокой разрешающей способности является очень ценным для реконструкции филогений (см. с. 4 работы [138]). Они опубликовали обширную сводку последовательностей кодонов для глобинов (на языке кода РНК), установленных ими с помощью этого метода. Поэтому было бы интересно проверить правильность полученных ими результатов, так как теперь благодаря новой методике амплификации генов в бактериальной плазмиде с после-

дующим быстрым секвенированием ДНК по методу Максама — Гилберта удалось установить нуклеотидные последовательности ряда глобиновых генов. Недавно, используя такую процедуру, Хейнделл и др. [177] расшифровали последовательность матричной РНК (мРНК) α -глобина кролика, которую можно сравнить с последовательностью кодонов, полученной Гудменом и др. [138] методом наибольшей экономии.

Оказывается, что из 51 кодона, предсказанного Гудменом и др. для α -глобина кролика (см. табл. 4 в их статье за 1974 г.), 44 просто неверны, а правильны всего 7 [242]. Другими словами, верными оказались только 14% «наиболее экономных» кодонов. Столь плохой результат поистине удивителен: большая часть ошибок приходится на третьи положения кодонов, между тем даже при случайном распределении четырех нуклеотидных оснований в третьих положениях кодонов могло бы получиться в среднем 25%-е совпадение. Таким же образом можно проверить адекватность «наиболее экономной» последовательности кодонов β -глобина человека (см. табл. 5 в работе [138]), поскольку Маротта и др. [328] уже определили полную нуклеотидную последовательность соответствующей мРНК. Здесь неверными оказываются около половины «наиболее экономных» кодонов, причем самая грубая ошибка заключается в том, что серину в 44-м положении аминокислотной последовательности, который в действительности кодируется триплетом UCC, поставлен в соответствие кодон AGU; таким образом, все три буквы этого кодона были предсказаны неверно.

Можно возразить, что нуклеотиды в третьем положении кодонов в многочисленных таблицах, составленных Гудменом и др. [138], не были указаны авторами совершенно однозначно. Как объясняется в несколько странном примечании к табл. 1 в работе [138], U в третьем положении обозначает или U, или C, а G — или A, или G; иногда же они могут обозначать любой из четырех нуклеотидов. В свете этого примечания приведенная выше критика «наиболее экономных» кодонов отчасти теряет смысл, что, однако, не может считаться свидетельством правильности «наиболее экономного» алгоритма Гудмена. «Наиболее экономные» кодоны действительно ошибочны, что показано весьма эффектным способом Холмквистом [182]. Этому исследованию предшествовали машинные эксперименты Татено и Нея [545], которые с помощью метода Монте-Карло моделировали процесс дивергенции нуклеотидных последовательностей за счет случайного накопления мутационных замен (каждая гипотетическая последовательность состояла из 300 нуклеотидов, соответствующих 100 кодонам, одна половина которых считалась вариабельной, а другая — постоянной). В результате было получено филогенетическое древо, отражающее эволюцию двадцати одной «современной» последовательности от общей

для них предковой последовательности А. Весьма существенно, что в этих модельных экспериментах в отличие от аналогичных исследований эволюционных процессов, протекающих в природе, известны точная топология, число накопленных нуклеотидных замен в каждой ветви филогенетического древа, а также все предковые последовательности в точках ветвления. Холмквист, получив от Нея и Татено результаты их имитационных экспериментов, послал двадцать одну «современную» последовательность (переведенную на язык аминокислотных последовательностей белков) Гудмену, обратившись к нему с просьбой дать «наиболее экономное» решение, рассматривающее как топологию, так и реконструированные предковые аминокислотные последовательности (что было названо решением с «предполагаемой топологией» — *inferred topology solution*). Холмквист просил также Гудмена представить ему реконструированные предковые аминокислотные последовательности, дающие правильную топологию (полученное таким образом решение было названо решением с «известной топологией» — *known topology solution*). Полученные Гудменом аминокислотные последовательности сравнивались потом с известными предковыми последовательностями. Оказалось, что в последовательностях, найденных методом наибольшей экономии, имеются многочисленные ошибки. Особое внимание обращает на себя высокая скорость накопления ошибок с увеличением расстояния от «узловых» последовательностей до «современных»: для узлов, удаленных на 200 или более замен, ошибка превышает 80%. В случае отдаленных предковых последовательностей, полученных в результате реконструкции, возникает (даже при учете правильной топологии) так много ошибок, что эти реконструированные последовательности не представляют никакой ценности (они даже содержат кодоны терминации синтеза полипептидных цепей!). Совершенно ясно, что любые рассуждения относительно ранних стадий эволюции глобинов, основанные на «наиболее экономных» реконструкциях, лишены смысла [243].

Все эти результаты четко показывают, что работы Гудмена и др. [138, 139] содержат слишком много ошибок и неточностей и, вопреки утверждениям авторов, никак не свидетельствуют об очень быстром накоплении аминокислотных замен на ранних этапах эволюции глобинов позвоночных. Выдвинутый ими тезис, согласно которому высокая скорость эволюции предковых глобинов связана с положительным дарвиновским отбором, несмотря на свое правдоподобие, не находит подтверждения. Как будет показано в гл. 5, гораздо более вероятно, что высокая скорость молекулярной эволюции обусловлена снятием функциональных ограничений, сопровождающимся превращением значительной части ранее вредных мутаций в селективно нейтральные, которые затем фиксируются в результате слу-

чайного дрейфа. Другими словами, необычайно высокая скорость эволюции той или иной молекулы свидетельствует не о приобретении ею какой-то новой функции, а скорее о снятии определенных функциональных ограничений.

Все это, конечно, не означает, что при эволюции глобинов не происходили те адаптивные мутационные замены, которые в ходе дифференциации, последовавшей за генной дупликацией, привели к возникновению функционирующего тетрамерного гемоглобина $\alpha_2\beta_2$ из его мономерного предшественника (в какой-то мере сходного с HbV миноги). В данном случае действительно должна была происходить длительная селективная «подгонка» субъединиц. Я же отстаиваю точку зрения, согласно которой на такие адаптивные замены, осуществлявшиеся при эволюции гемоглобина, наложилось значительно больше селективно нейтральных или почти нейтральных замен, зафиксировавшихся в процессе случайного дрейфа.

Сравнительное изучение нуклеотидных последовательностей

Хотя в течение последних примерно 15 лет изучение молекулярной эволюции основывалось главным образом на результатах сравнительных исследований аминокислотных последовательностей белков (рассматривавшихся в предыдущем разделе), в настоящее время с появлением огромного количества данных о нуклеотидных последовательностях все большее значение приобретают именно эти материалы. Их получение стало возможным благодаря разработке принципиально новой методики амплификации генов в бактериальных плаزمидях с последующим быстрым секвенированием ДНК по методу Максама — Гилберта [355]. В последние годы объем данных о нуклеотидных последовательностях различных участков геномов самых разных организмов растет столь стремительно, что в скором времени сравнительные исследования аминокислотных последовательностей, по-видимому, будут полностью «вытеснены» сравнительными исследованиями нуклеотидных последовательностей.

Становится все более очевидным, что преобладание синонимичных и других молчащих замен оснований — удивительная общая особенность молекулярной эволюции и что этот факт согласуется с теорией нейтральности [202, 204, 205, 210, 236, 239, 241, 244, 363, 365, 422].

При расчете эволюционных расстояний между гомологичными полинуклеотидами, выраженных через число нуклеотидных замен, необходимо вводить поправки на множественные и обратные замены в гомологичных сайтах. Это связано с тем, что полинуклеотиды состоят из оснований всего четырех типов, и

поэтому даже две случайные последовательности совпадают в среднем по 25% сайтов. Таким образом, нужны адекватные статистические методы оценки эволюционных расстояний (а также эволюционных скоростей, когда нам известно время дивергенции), позволяющие вводить необходимые поправки.

Рассмотрим пару гомологичных сайтов двух нуклеотидных последовательностей (называемых ниже последовательностями

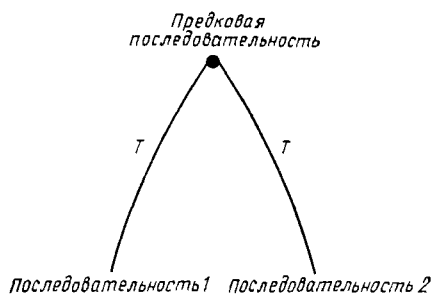


Рис. 4.8. Дивергенция двух гомологичных последовательностей от общей для них предковой последовательности, существовавшей T лет назад.

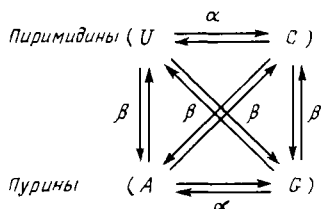


Рис. 4.9. Схема эволюционных замен оснований. В рамках этой модели скорость транзиций α может отличаться от скорости трансверсий β .

ми 1 и 2). В дальнейшем мы будем использовать язык последовательности РНК (а не ДНК), т. е. последовательности, состоящей из четырех азотистых оснований U, C, A и G. Это облегчит переход от кодонов (триплетов) к аминокислотам (и обратно) с помощью таблицы стандартного кода. Проанализируем дифференциацию последовательностей 1 и 2 по данным гомологичным сайтам в ходе эволюции от общей для них предковой последовательности, существовавшей T лет назад (рис. 4.8). В этих сайтах в ходе эволюции происходят последовательные замены оснований. Для формального описания процесса воспользуемся моделью эволюционных замен оснований, представленной на рис. 4.9. Здесь α — скорость замен по типу транзиции, а 2β — скорость замен по типу трансверсии; таким образом, общая скорость замен на сайт в год равна $k = \alpha + 2\beta$. Заметим, что α и β — это не обычные скорости мутирования на уровне особей, а эволюционные скорости мутационных замен на уровне видов.

Общее число замен оснований, накопленных в двух ветвях, длина каждой из которых равна T (см. рис. 4.8), есть $2kT$:

$$K = 2Tk = 2\alpha T + 4\beta T. \quad (4.9)$$

Таблица 4.7. Различные комбинации оснований, находящихся в гомологичных нуклеотидных сайтах двух последовательностей

	Тип различия		
	транзиции	трансверсии	
Последовательность 1	U C A G	U A C G	U G C A
Последовательность 2	C U G A	A U G C	G U A C
(Частота)	(P)	(Q)	

Сравнивая гомологичные сайты последовательностей 1 и 2, можно видеть, что существует двенадцать комбинаций разных оснований (см. табл. 4.7).

Пусть P — вероятность (относительная частота) того, что в момент времени T гомологичные сайты заняты основаниями U и C, C и U, A и G, G и A. Иными словами, P — это вероятность того, что гомологичные сайты различаются по типу транзиции. Аналогичным образом пусть Q — вероятность того, что гомологичные сайты различаются по типу трансверсии (см. табл. 4.7). Можно показать, что P и Q удовлетворяют следующей системе дифференциальных уравнений (вывод которых см. в работе [241]):

$$\frac{dP}{dT} = 2\alpha - 4(\alpha + \beta)P - 2(\alpha - \beta)Q, \quad (4.10)$$

$$\frac{dQ}{dT} = 4\beta - 8\beta Q.$$

Решение этой системы, удовлетворяющее условию $P=Q=0$ при $T=0$, т. е. условию исходной идентичности двух последовательностей, имеет следующий вид:

$$P = \frac{1}{4} - \frac{1}{2} e^{-4(\alpha + \beta)T} + \frac{1}{4} e^{-8\beta T}, \quad (4.11)$$

$$Q = \frac{1}{2} - \frac{1}{2} e^{-8\beta T}.$$

Отсюда

$$4(\alpha + \beta)T = -\ln(1 - 2P - Q), \quad (4.12)$$

и

$$8\beta T = -\ln(1 - 2Q). \quad (4.13)$$

Так как эволюционное расстояние, выраженное через число замен оснований, по которым различаются две нуклеотидные последовательности, задается равенством (4.9), мы получаем

$$K = -\frac{1}{2} \ln [(1 - 2P - Q) \sqrt{1 - 2Q}] \quad (4.14)$$

[241]. Приведенная формула может использоваться для оценки эволюционного расстояния между двумя последовательностями, выраженного через число замен оснований, происшедших за T лет эволюции в расчете на сайт. Здесь $P = n_1/n$, а $Q = n_2/n$, где n_1 и n_2 — числа сайтов двух сравниваемых последовательностей, различающихся заменами по типу транзиции и трансверсии соответственно, а n — общее число сравниваемых сайтов.

В частном случае $\alpha = \beta$ выражения (4.11) сводятся к равенству $P = Q/2 = [1 - \exp(-8\alpha T)]/4$, и, следовательно, (4.14) принимает вид

$$K = -\frac{3}{4} \ln \left(1 - \frac{4}{3} \lambda \right), \quad (4.15)$$

где λ — доля сайтов, по которым различаются две сравниваемые последовательности. Эта формула хорошо известна (см. [258]) и впервые была получена Джуксом и Кантором [206]. Однако на самом деле — особенно при сравнении по третьим положениям кодонов — P часто бывает больше Q , и, следовательно допущение, что $\alpha = \beta$, а $P = Q/2$, не всегда реалистично. Это одна из причин, по которой формула (4.14) предпочтительнее формулы (4.15). Равенство (4.14) обладает также одним ценным свойством: при малых P и Q оно сводится — независимо от значений α и β — к выражению $K = P + Q$. В этом случае, зная время дивергенции T , можно найти скорость замен оснований в год: $k_{\text{нукл}} = K/(2T)$, где индекс «нукл» означает, что эта скорость оценивается в расчете на нуклеотидный сайт.

Так как значительная часть замен по третьим положениям кодонов синонимична, т. е. не приводит к заменам аминокислот, было бы интересно оценить синонимичную компоненту скорости замен по этому положению. Из стандартной таблицы кода РНК видно, что для любой данной пары оснований в первом и втором положениях кодона возможны, вообще говоря, две ситуации: либо по третьему положению кодона имеет место полная синонимичность (четырёхкратная вырожденность), либо синонимичность ограничена пуринами или пиримидинами (двукратная вырожденность). Эти две ситуации встречаются с примерно одинаковой частотой. Поэтому синонимичную компоненту скорости замен по третьему положению кодонов, которую мы обозначим через k_s' , можно оценить по формуле $k_s' = 0,5(\alpha + 2\beta) + 0,5\alpha$, или $k_s = \alpha + \beta$. Пусть $K_s' = 2Tk_s' = 2(\alpha + \beta)T -$

синонимичная компонента расстояния; тогда, используя равенство (4.12), получаем

$$K_S' = -\frac{1}{2} \ln(1 - 2P - Q). \quad (4.16)$$

Имеет смысл найти формулы для дисперсии оценок K и K_S' , обусловленной ошибкой выборки. Пусть n — число нуклеотидных сайтов, по которым сравниваются две последовательности. Тогда стандартные ошибки K и K_S' будут равны

$$\sigma_K = \frac{1}{\sqrt{n}} \sqrt{[(a^2P + b^2Q) - (aP + bQ)^2]}, \quad (4.17)$$

$$\sigma_{K_S'} = \frac{\sqrt{[4P + Q - (2P + Q)^2]}}{2(1 - 2P - Q)\sqrt{n}}, \quad (4.18)$$

где $a = 1/(1 - 2P - Q)$, $b = (1/2) \cdot [1/(1 - 2P - Q) + 1/(1 - 2Q)]$.

Для примера сравним нуклеотидные последовательности генов β -цепей глобинов кролика [92] и курицы [482]. Эти последовательности можно сравнивать по 438 нуклеотидным сайтам, соответствующим 146 аминокислотным сайтам (кодонам). В 58 сайтах этих двух последовательностей обнаруживаются различия по типу транзиции, а в 63 — по типу трансверсии. Таким образом, $P = 0,132$, $Q = 0,144$, и мы получаем $K = 0,348$. Поскольку млекопитающие и птицы дивергировали, по всей вероятности, в каменноугольном периоде, мы можем в качестве ориентировочного значения принять $T = 300 \cdot 10^6$ годам. Тогда скорость эволюции этих последовательностей в расчете на сайт будет равна $k_{\text{нукл}} = K/(2T) = 0,58 \cdot 10^{-9}$ в год. Эта величина представляет собой общую скорость замен в расчете на сайт, однако гораздо интереснее оценить скорость эволюции для каждого из трех положений кодонов. Для первого положения, по которому, как и по двум другим, сравнивалось 146 сайтов, $P = 15/146$, $Q = 21/146$, что дает $K_1 = 0,300$ (индекс 1 указывает, что значение K относится к первому положению кодонов). Аналогичным образом для второго положения имеем $P = 7/146$, $Q = 18/146$, так что $K_2 = 0,195$. И наконец, для третьего положения $P = 36/146$, $Q = 24/146$; теперь $K_3 = 0,635$, что намного больше, чем в первых двух случаях. Можно также оценить синонимичную компоненту эволюционного расстояния по третьему положению. Вычисление ее с помощью формулы (4.16) дает $K_S' = 0,535$.

В табл. 4.8 приведены результаты аналогичных расчетов для других последовательностей. Помимо генов β -цепей глобинов курицы и кролика рассматриваются гены β -цепей глобина человека [328] и мыши [283], а также α -цепей глобина кролика [177]. Из этой таблицы видно, что для всех рассмотренных случаев, за исключением двух последних, в которых проводит-

Таблица 4.8. Эволюционные расстояния, выраженные через число замен оснований и рассчитанные по результатам сравнений нуклеотидных последовательностей генов нескольких глобинов; K_1 , K_2 и K_3 — число замен оснований в первом, втором и третьем положениях кодонов, а K'_s — расчетное число синонимичных замен в третьем положении. Приведены расчетные значения этих параметров и их стандартные ошибки

	Эволюционные расстояния в расчете на нуклеотидный сайт			
	K_1	K_2	K_3	K'_s
β -цепи курицы и кролика	$0,30 \pm 0,05$	$0,19 \pm 0,04$	$0,64 \pm 0,11$	$0,53 \pm 0,10$
β -цепи человека и мыши	$0,17 \pm 0,04$	$0,13 \pm 0,03$	$0,34 \pm 0,06$	$0,28 \pm 0,05$
β -цепи человека и кролика	$0,06 \pm 0,02$	$0,06 \pm 0,02$	$0,28 \pm 0,06$	$0,25 \pm 0,05$
β -цепи кролика и мыши	$0,16 \pm 0,04$	$0,13 \pm 0,03$	$0,43 \pm 0,07$	$0,36 \pm 0,07$
α -цепь кролика и β -цепь кролика	$0,54 \pm 0,09$	$0,44 \pm 0,07$	$0,90 \pm 0,15$	$0,69 \pm 0,13$
α -цепь кролика и α -1-цепь мыши	$0,12 \pm 0,03$	$0,11 \pm 0,03$	$0,54 \pm 0,09$	$0,47 \pm 0,09$
α -цепь кролика и α -3-цепь мыши	$0,27 \pm 0,06$	$0,28 \pm 0,06$	$0,69 \pm 0,13$	$0,56 \pm 0,12$
α -1-цепь мыши и α -3-цепь мыши	$0,16 \pm 0,04$	$0,20 \pm 0,05$	$0,30 \pm 0,06$	$0,22 \pm 0,05$

ся сравнение с аномальным глобиноподобным α -3-геном [421], как правило, справедливо соотношение $K_2 < K_1 < K_3$, т. е. наибольшая скорость мутационных замен характерна для третьего положения кодонов, меньшая — для первого и самая маленькая — для второго. Это легко объяснить в рамках теории нейтральности, так как замены нуклеотидов во втором положении обычно сопровождаются заменой на такие аминокислоты, физико-химические свойства которых отличаются от свойств замещаемых аминокислот значительно сильнее, чем в случае замены по первому и третьему положениям. Рассмотрим, например, кодон для Pго. Согласно таблице генетического кода (см. табл. 2.1), этим кодоном является CCN (N — любое из четырех оснований). Замены основания C в первом положении на U, A и G приводят к заменам Pго на Ser, Thr и Ala соответственно. Как показывает расчет расстояний, основанный на различиях аминокислот по полярности и объему (см. [362]), эти аминокислоты удалены от Pго на 0,56, 0,87 и 0,06 единиц, а среднее расстояние между ними и Pго равно примерно 0,5. В то же время в случае замены C во втором положении кодона соответ-

ствующее среднее расстояние оказывается равным примерно 2,5.

Это означает, что мутационные замены в первом положении с большей вероятностью могут оказаться селективно нейтральными (т. е. селективно эквивалентными), чем во втором, и поэтому они с большей вероятностью фиксируются в результате случайного дрейфа [261]. Такие же соображения с еще большим основанием применимы при сравнении замен в третьем положении с заменами в первом и втором положениях, так как большинство замен в третьем положении вообще не сопровождается заменами аминокислот.

В качестве показателя относительных частот синонимичных и несинонимичных замен можно использовать отношение K_s' : $\frac{1}{2} (K_1 + K_2)$. В случае β -цепей гемоглобина человека и кролика ($T = 8 \cdot 10^7$ лет) это отношение равно 4,17, β -цепей гемоглобина курицы и кролика ($T = 3 \cdot 10^8$ лет) — 2,16, а α - и β -цепей гемоглобина кролика ($T = 3 \cdot 10^8$) — всего 1,41. Перье и др. [466], основываясь на аналогичном исследовании эволюции предшественника проинсулина и глобина, пришли к заключению, что на ранних этапах эволюции этих белков в «молчащих» сайтах замены накапливались примерно в 7 раз быстрее, чем в сайтах, замены по которым приводят к замещению аминокислот, однако этот процесс накопления достиг насыщения уже 85—100 млн. лет назад. Такой феномен насыщения, вероятно, представляет собой артефакт, обусловленный трудностью выявления всех синонимичных замен по мере их накопления в кодонах. Одна из причин, побудивших меня выдвинуть это предположение, заключается в том, что при сравнении менее отдаленных друг от друга последовательностей (например, генов млекопитающих) были получены более высокие значения скоростей эволюции; так как эволюция протекала в направлении от прошлого к настоящему, а не наоборот, феномен насыщения может быть только кажущимся. Если бы темпы эволюционных замен действительно снижались, при сравнении менее «отдаленных» последовательностей должно было бы наблюдаться более медленное накопление замен, однако на самом деле ситуация здесь прямо противоположная. Кроме того, для третьего положения кодонов — по сравнению с первым и вторым — имеет место четкое отклонение от правила Чаргаффа (так, в гене β -цепи глобина человека содержание G составляет 36%, C — 30%, U — 27% и A — 7%), и это затрудняет получение оценок в случае очень больших значений T . С другими аспектами проблемы определения числа синонимичных замен в третьем положении кодонов читатель может ознакомиться в одной из статей Кимуры [244].

Гораздо более интересными представляются два последних случая, рассмотренных в табл. 4.8, где проводится сравнение

с «глобиноподобным» α -3-геном мыши, у которого недавно была определена нуклеотидная последовательность [421]. В этом гене нет двух вставных последовательностей, в норме присутствующих в генах всех α - и β -глобинов, и он не кодирует никакого глобина. Иными словами, этот ген неактивен в том смысле, что на нем не синтезируется стабильная мРНК. Однако при сравнении этой последовательности с нуклеотидными последовательностями генов α -глобина (α -1) взрослой мыши и α -глобина кролика становится очевидно, что α -3-ген произошел от нормального предкового гена α -глобина в результате его дупликации и последующей потери вставок. Это должно было произойти после дивергенции мыши и кролика от их общего предка, происшедшей около 80 млн. лет назад. В кодирующем участке рассматриваемого глобиноподобного α -гена произошло много вставок и делеций. Поэтому при сравнении нуклеотидных последовательностей я рассматривал только те кодоны α -3-гена, которые не затронуты такими структурными перестройками и которые или идентичны соответствующим (гомологичным) кодонам α -1-гена мыши и α -гена кролика, или отличаются от них лишь заменами оснований.

Как видно из трех последних строк табл. 4.8, этот «мертвый ген» (α -3) эволюционировал с гораздо большей скоростью, чем его нормальный двойник (α -1-ген мыши), особенно по первому и второму положениям кодонов. Это легко объяснить в рамках теории нейтральности. В нормальных условиях для каждого гена существуют селективные ограничения, определяемые требованиями, предъявляемыми к продуцируемому им белку, который должен быть функционально активным, поэтому эволюционные изменения ограничиваются строго определенным набором замен оснований. Однако, когда ген освобождается от этих ограничений, как в случае глобиноподобного α -3-гена, практически любая замена основания в нем уже не оказывает влияния на «дарвиновскую» приспособленность и скорость замен оснований должна приблизиться к своему верхнему пределу, определяемому темпами мутирования (это произойдет только в том случае, если верна теория нейтральности, и не будет иметь места, если большинство замен оснований обусловлено действием положительного отбора). Если скорости синонимичных замен не очень далеки от этого предела [236], можно ожидать, что скорости эволюции «мертвых генов» примерно равны скоростям синонимичных замен. Недавно Мията и Ясунага [364] рассчитали скорость эволюции псевдоальфаглобинового гена ($\psi\alpha$ 30.5) мыши, нуклеотидная последовательность которого была определена Ваниным и др. [557] и который, по-видимому, эквивалентен α -3-гену. Они тоже получили результат, подтверждающий высказанное предположение.

Было бы интересно оценить скорости замен оснований не

только в кодирующих участках, но и в некодирующих, например в интронах. Для этого я использовал данные, представленные ван Ойеном и др. [558], которые занимались сравнительным анализом нуклеотидных последовательностей генов β -цепей глобинов кролика и мыши. В табл. 2 их работы приведены данные о числе транзиций и трансверсий (отдельно), по которым различаются гомологичные участки этих последовательностей. Для небольших интронов, исключая пять «пробелов» суммарной длины до шести нуклеотидов, $P=27/113$, $Q=18/113$ и $n=113$. Используя формулы (4.14) и (4.17), мы получаем $K=0,60\pm0,12$. Эта величина не очень отличается от скорости замен в третьем положении $K_3=0,43\pm0,07$. Большие интроны генов β -цепей глобинов кролика и мыши имеют разную длину и отличаются друг от друга четырнадцатью пробелами (выявленными путем оптимизации соответствия их последовательностей), составляющими в сумме 109 нуклеотидов. Если исключить эти участки, то окажется, что $\bar{P}=113/557$, $Q=179/557$ и $K=0,90\pm0,07$. Эта величина достоверно выше приведенного значения K_3 . По-видимому, как отмечали ван Ойен и др. [558], в этом участке гена помимо точковых мутаций довольно часто происходят вставки и делеции, которые сильно увеличивают расчетное значение «скорости нуклеотидных замен», так как большинство таких перестроек тоже могут быть селективно нейтральными и подвергаться случайной фиксации в процессе генетического дрейфа.

Некоторые особенности молекулярной эволюции

Молекулярная эволюция имеет ряд особенностей, отличающих ее от эволюции фенотипической. Эти особенности можно представить в виде пяти принципов [261], которые мы и рассмотрим в настоящей главе. Первые четыре из них — эмпирические, а пятый, помогающий понять природу «селективно нейтральных мутаций», установлен теоретическим путем.

1. Скорость эволюции любого белка, выраженная через число аминокислотных замен на сайт в год, приблизительно постоянна и одинакова в разных филогенетических линиях, если только функция и третичная структура этого белка остаются в основном неизменными.

Здесь сформулирована гипотеза о постоянстве скорости молекулярной эволюции, которую Цукеркандл и Полинг [622] назвали концепцией «молекулярных эволюционных часов». Проблема эволюционных часов принадлежит к числу наиболее дискуссионных аспектов молекулярной эволюции. Особенно оживленное ее обсуждение развернулось в связи с датировкой времени дивергенции человека и человекообразных обезьян (см. [584]) и появлением данных как за, так и против теории нейтральности. Проблемы, связанные с постоянством темпов молекулярной эволюции, подробно рассмотрены в гл. 4, и я не буду останавливаться на них еще раз. Напомню только, что наиболее убедительные доказательства справедливости концепции молекулярных часов были получены при изучении глобинов млекопитающих, хотя наблюдаемая дисперсия темпов эволюции в разных филогенетических линиях оказалась несколько больше ожидаемой из чисто вероятностных соображений. Приблизительное постоянство темпов эволюции характерно также для цитохрома *c* и фибринопептидов [89], хотя скорость эволюции каждого из этих белков отличается от скорости эволюции глобинов (см. также табл. 4.1). Недавно Дэйхоф [77] опубликовала обширный список «скоростей включения мутаций» («rates of mutation acceptance»), т. е. скоростей молекулярной эволюции, выраженных через число мутационных за-

мен, для различных семейств белков, для которых можно получить достоверные оценки. Она пришла к выводу, что, хотя скорость изменения наиболее быстро эволюционирующих белков почти в 400 раз больше, чем у белка с самой низкой скоростью эволюции — гистона, темпы эволюции белков, принадлежащих к одному семейству, редко различаются более чем в два или три раза, особенно если эти белки занимают одинаковые «функциональные ниши» у разных организмов. Тот факт, что филогенетические древа, построенные при сравнительном изучении аминокислотных и нуклеотидных последовательностей, согласуются с обычными филогенетическими схемами, по-видимому, свидетельствует о постоянстве темпов молекулярной эволюции.

Сказанное не означает, что концепция молекулярных часов абсолютно правильна. В некоторых случаях она действительно кажется не совсем согласующейся с фактами. Например, по данным работы Матсубары и др. [345], эволюция ферредоксина хлоропластов растений у покрытосеменных протекала гораздо быстрее, чем у сине-зеленых водорослей. Однако эти авторы были довольно осторожны в своих выводах, отметив, что время ответвления линии покрытосеменных пока неизвестно и что полученные ими данные могут свидетельствовать о полифилетическом происхождении цветковых растений. Джукс и Холмквист [207] сообщили, что цитохром *c* гремучей змеи эволюционировал в 3—4 раза быстрее цитохрома *c* черепахи. Более систематическое исследование проблемы постоянства темпов молекулярной эволюции недавно провел Снит [521], который указал на необходимость отличать скорость наблюдаемых замен от «истинной скорости» („implied underlying rate”). Снит рассчитал коэффициенты вариации истинных скоростей, используя данные о цитохроме *c*, глобинах, фибринопептидах, панкреатической рибонуклеазе, токсине змеиного яда, и получил результаты, свидетельствующие о замечательном постоянстве истинных скоростей эволюции всех этих белков, за исключением цитохрома *c*. Правда, Снит делает оговорку, что при оценке таких эволюционных скоростей используются не совсем обоснованные допущения.

Уилсон и др. [584] попытались проверить правильность гипотезы о зависимости темпов молекулярной эволюции от продолжительности поколения, т. е. предположения о том, что накопление замен зависит в первую очередь от числа поколений, сменившихся с момента дивергенции, и лишь во вторую очередь от соответствующего абсолютного времени. Используя данные о большом числе макромолекул, они сравнили скорости их эволюции у млекопитающих с малой и большой продолжительностью поколения. Оказалось, что ни одна из наблюдаемых величин не попадает в интервал, предсказанный на основе этой гипотезы. С другой стороны, наблюдаемые величины

согласуются с предположением о прямо пропорциональной зависимости между накоплением замен и абсолютным временем. Возможность корректной экстраполяции темпов молекулярной эволюции высших животных на бактерий, основанной на гипотезе о постоянстве скорости, представляется спорной и даже сомнительной. Однако очевидный успех в построении по данным о рибосомных 5S-РНК [185] эволюционного древа для всех царств живых организмов, согласующегося с обычным родословным древом, свидетельствует о том, что даже в этом случае гипотеза молекулярных часов не может быть полностью отвергнута. То, что скорость эволюции, выраженная через число мутационных замен в определенной макромолекуле (скажем, в α -цепи гемоглобина млекопитающих) в расчете на год постоянна и совсем не зависит от условий, образа жизни и, по-видимому, продолжительности поколения данного вида, гораздо легче объяснить на основе теории нейтральности, чем с позиций селекционизма. Абсолютное постоянство темпов молекулярной эволюции не является обязательным условием правильности теории нейтральности, хотя, если она действительно верна, все исключения должны быть ею объяснены.

Как показано в гл. 3, согласно теории нейтральности $k = v$ [формула (3.20)], т. е. для нейтральных аллелей скорость эволюции равна скорости мутирования. С учетом того, что определенная доля мутаций может быть вредной, это равенство можно также представить в виде

$$k = v_0 = v_T f_0, \quad (5.1)$$

где v_T — полная скорость мутирования, f_0 — доля нейтральных мутаций. Обратите внимание, что существование вредных мутаций не противоречит теории нейтральности, так как они не вносят вклада в имеющиеся эволюционные изменения. Чтобы объяснить факты, следует предположить, что для любой данной молекулы скорость появления нейтральных мутаций в расчете на год постоянна. Конечно, могут возникнуть и благоприятные мутации, однако мы предполагаем, что они слишком редки, чтобы оказывать какое-либо влияние на темпы молекулярной эволюции.

С другой стороны, если динамика мутационных замен определяется положительным естественным отбором, то $k = 4N_e s v$ [формула (3.22)], т. е. скорость молекулярной эволюции равна учетверенному произведению эффективного размера популяции (N_e) на скорость мутирования (v) к аллелям, обладающим явным преимуществом, и на коэффициент отбора (s). Если учесть еще возможность возникновения вредных мутаций, «селекционистская» формула, отвечающая (5.1), примет вид

$$k = 4N_e \bar{s}_A v_T f_A, \quad (5.2)$$

где v_t — полная скорость мутирования, f_A — доля благоприятных мутаций, $\bar{s}_A$ — средний коэффициент отбора таких мутаций. Отметим, что произведение $N_e \bar{s}_A$ должно быть много больше единицы, в противном случае динамика этих мутаций будет больше похожа на динамику селективно нейтральных, чем селективно благоприятных мутаций. Здесь мы пренебрегаем нейтральными или почти нейтральными мутациями, так как с позиций селекционистской теории они не вносят существенного вклада в мутационные замены. В таком случае для объяснения наблюдаемого постоянства скоростей молекулярной эволюции надо принять предположение о постоянстве и равенстве произведения $N_e \bar{s}_A v_t f_A$ для разных родословных любого данного белка. Иными словами, для объяснения приблизительного постоянства темпов молекулярной эволюции в расчете на год с селекционистских позиций необходимо принять ряд взаимосвязанных и весьма произвольных предположений относительно частоты мутирования, взаимодействия генов, экологических условий и численности популяции. С другой стороны, гипотеза нейтральности требует принятия только одного допущения *ad hoc*, а именно предположения об увеличении f_0 с увеличением продолжительности поколения. Это предположение представляется достаточно реалистическим, если учесть, что более крупные организмы с более продолжительным жизненным циклом, как правило, образуют популяции меньшего размера, и, следовательно, доля селективно нейтральных мутаций будет у них больше. Мы рассмотрим этот вопрос более детально в разд. 8.7. Следует отметить, что при сравнении двух организмов, имеющих одинаковую продолжительность поколения, но обитающих в совершенно разных условиях, такой проблемы не возникает.

Один из важных вопросов, нередко упускаемых из виду, касается влияния численности популяции; часто утверждается (см., например, [466], с. 561), что постоянство эволюционной скорости фиксации мутаций можно объяснить постоянством «коэффициента отбора». Такое утверждение, очевидно, ошибочно. Если эволюция — результат постепенного накопления редких благоприятных мутаций, то в больших популяциях такие мутации должны были бы появляться в значительно большем количестве, чем в малых, и поэтому в первых молекулярная эволюция протекала бы намного быстрее, чем во вторых.

В отличие от фенотипической эволюции, генетические механизмы которой в деталях пока не известны, в случае молекулярной эволюции можно провести гораздо более глубокое исследование процесса генетических изменений, лежащих в ее основе.

Хорошо известно, что появление большинства аминокислот-

ных замен в ходе эволюции, а также вариантов гемоглобина, обнаруженных в популяциях человека, можно объяснить единичными заменами оснований (см., например, [561]). В глобинах, как мы уже знаем, аминокислотные замены происходят со скоростью примерно 10^{-9} на аминокислотный сайт в год. Это означает, что, скажем, в α -цепи гемоглобина, состоящей из 141 аминокислоты, одна замена происходит в среднем каждые 7 млн. лет. Рассмотрим некий вид млекопитающих со средней продолжительностью поколения (g), равной двум годам, и предположим, что в каждом поколении насчитывается полмиллиона особей (при этом эффективный размер может быть намного меньше). Пусть скорость мутирования (v) гена α -цепи глобина в расчете на поколение равна 10^{-6} . Ограничим наше рассмотрение только теми мутациями, которые вызывают замены аминокислот. Полное число мутаций в гене α -цепи, происшедших у данного вида за 7 млн. лет, будет тогда равно $2Nv \cdot 7 \cdot 10^6 / g = 2 \cdot (0,5 \cdot 10^6) \cdot 10^{-6} \cdot 7 \cdot 10^6 / 2 = 3,5 \cdot 10^6$. Таким образом, из 3,5 млн. возникших мутаций зафиксорируется только одна, причем не все эти мутации будут уникальны.

Оценим теперь число одинаковых мутаций, повторно возникших за указанный промежуток времени. Для этого рассмотрим вопрос об общем числе аллелей данного локуса, которые могут возникнуть вследствие замен отдельных оснований. Хотя каждый аминокислотный сайт (кодон) может быть занят одной из 20 аминокислот, число аминокислот, которые могут быть порождены заменой отдельных оснований, в расчете на кодон оказывается меньше. Среднее их число равно примерно 6,5. Это можно показать следующим образом. Рассмотрим замены оснований в отдельных нуклеотидных сайтах кодона. Замены оснований в первых и во вторых положениях, как правило, приводят к разным аминокислотам, и поэтому возможное число разных аминокислотных замещений в этом случае равно $3+3$, или 6. Для третьего положения существуют, вообще говоря, две возможности, реализующиеся с одинаковыми частотами. Если кодон характеризуется четырехкратной вырожденностью (как в случае Val), то замена основания в третьем положении кодона не приводит к замене аминокислоты; при двукратной же вырожденности такая замена возможна (как, например, в случае Asp и Glu, кодируемых триплетом GAN). Поэтому число различных аминокислотных замен, вызываемых заменами в третьем положении кодонов, в среднем равно 0,5. Таким образом, полное число разных аминокислотных замен, обусловленных заменами отдельных оснований, в расчете на кодон равно $3+3+0,5$, или 6,5. Для всей α -цепи глобина приближенное число возможных аллелей, которые могут образоваться в результате замены отдельных оснований, равно $6,5 \cdot 141$, или 916,5.

Продолжая наш приближенный расчет, примем, что число α -глобиновых аллелей равно 1000; тогда, разделив 3,5 млн. мутаций на 1000, получим, что за 7 млн. лет одинаковые мутантные аллели возникают 3500 раз. Если бы какой-нибудь из этих 1000 аллелей обладал значительным селективным преимуществом, то он зафиксировался бы у данного вида без повторных мутаций. Как показано в гл. 3, если мутация имеет селективное преимущество s , вероятность ее фиксации в предположении, что $4N_e s \gg 1$, равна $u = 2sN_e/N$ [см. формулу (3.19)]. Если эффективная численность составляет половину фактической ($N_e = N/2$), то $u = s$. При этом вероятность потери всех мутаций данного аллеля, за исключением 3500-й, равна e^{-3499s} . Эта вероятность должна быть довольно высокой, поскольку именно такая ситуация характерна для эволюции глобинов.

Примем эту вероятность равной 0,5, т. е. будем считать, что $e^{-3499s} = 0,5$. Отсюда $s = 2 \cdot 10^{-4}$. Если число различных благоприятных аллелей равно m , то приведенную величину s следует разделить на m . Аналогичную оценку s_A можно получить, если приравнять величину, обратную $k = 4N_e v_T f_A s_A$, к $7 \cdot 10^6$ и принять $4N_e = 10^6$, $v_T = 10^{-6}$ и $f_A = m/1000$. Столь малое селективное преимущество, даже если оно существует, никогда не удастся измерить на практике. В то же время с точки зрения теории нейтральности все выглядит намного проще; скорость мутирования к нейтральным аллелям в расчете на поколение равна $141 \cdot 10^{-9}$, или $v_0 = 1,4 \cdot 10^{-7}$, что составляет примерно 1/7 от предполагаемой полной скорости мутирования $v_T = 10^{-6}$; остальные мутации являются вредными.

Прежде чем переходить к обсуждению следующего принципа молекулярной эволюции, я хотел бы кратко рассмотреть тезис, согласно которому гемоглобины различных млекопитающих не являются физиологически эквивалентными, причем различия между ними отражают приспособление животных к среде обитания. Например, по данным Томиты и Ригса [551], для гемоглобина мыши характерен наибольший эффект Бора, для гемоглобина слона — наименьший, а гемоглобин человека занимает промежуточное положение между ними (хотя его отличие от гемоглобина мыши, по-видимому, очень невелико). Эффект Бора состоит в том, что при увеличении парциального давления углекислого газа или при понижении pH молекула гемоглобина стремится высвободить кислород, даже если его парциальное давление остается прежним. Предполагается, что, чем мельче млекопитающее, тем интенсивнее его метаболизм. Поэтому авторы утверждают, что выявленная зависимость между массой тела животного и эффектом Бора должна иметь определенное адаптивное значение и что это противоречит теории нейтральности. Я не могу опровергнуть данный тезис, однако

думаю, что адаптивные замены, если они вообще происходят, вероятно, составляют небольшую долю от общего числа аминокислотных замен, осуществляемых в ходе эволюции. В настоящее время существует по крайней мере несколько тысяч видов млекопитающих. Довольно трудно поверить, что каждая из молекул глобинов, специфичных для разных видов, обладает особенностями, обеспечивающими ее точное приспособление к потребностям какого-то определенного вида, и что эти потребности должны поэтому отличаться от точно таких же связанных с транспортом кислорода в крови потребностей у всех прочих видов.

2. Функционально менее важные молекулы или их части эволюционируют (накапливая мутационные замены) быстрее, чем более важные.

3. Мутационные замены, приводящие к меньшим нарушениям структуры и функции молекулы (консервативные замены), в ходе эволюции происходят чаще тех, которые вызывают более существенное нарушение структуры и функции этой молекулы.

Эти две закономерности очень важны для теории нейтральной молекулярной эволюции, и поэтому в дальнейшем, в гл. 7, мы рассмотрим их подробно. Здесь я только отмечу, что с позиций теории нейтральности они объясняются очень просто: вероятность появления мутации, не являющейся вредной, а точнее — вероятностью селективно нейтральной мутации, будет больше, если эта мутация возникает в функционально менее важной молекуле или ее части, так как она имеет большие шансы зафиксироваться в популяции в результате случайного генетического дрейфа. Эта вероятность также будет выше в случае «консервативных» мутаций, т. е. мутаций, не приводящих к коренным изменениям соответствующих макромолекул. Один из самых крайних случаев — синонимичные замены, происходящие преимущественно в третьем положении кодонов.

4. Появлению нового в функциональном отношении гена всегда должна предшествовать дупликация гена.

Важность дупликаций генов для эволюции организмов отмечалась еще выдающимися исследователями дрозофилы из школы Моргана. Здесь следует напомнить ключевой тезис, согласно которому благодаря существованию двух копий данного гена в одной из них могут накапливаться мутации и в конечном счете она превратится в новый ген, тогда как другая копия сохраняет старую функцию, необходимую виду для выживания в переходный период. «Защищенные» нормальной копией гена мутации, которые до дупликации были бы забракованы отбо-

ром, теперь накапливаются, а их накопление может привести к появлению нового гена.

Большой вклад в осознание «творческой» роли дубликации генов в эволюции организмов внесла очень интересная книга Оно [429], где рассмотрены новые доказательства эволюционного значения дупликаций, основанные на результатах современных молекулярно-биологических, цитологических и палеонтологических исследований. Несколько раньше Оно предложил смелую гипотезу, согласно которой наши позвоночные предки на стадии рыб, до выхода на сушу, вероятно, прошли через две тетраплоидизации [432]. Это произошло до установления жесткого хромосомного механизма определения пола. Упомянутая работа Оно, а также его более поздние исследования [431] внесли значительный вклад в развитие современной эволюционной теории, обратив внимание на удивительно консервативную природу мутационных замен. Большое сходство X-хромосом млекопитающих («правило Оно») легко объяснить, если учесть, что эти хромосомы находятся в единичной дозе как у самцов, так и фактически у самок (вследствие инактивации одной из них) и поэтому подвергаются более жесткому естественному отбору, чем аутосомы.

За последние несколько лет накопилось много данных, свидетельствующих о широкой распространенности дублицированных или повторяющихся структур в геномах высших организмов. Из предыдущей главы мы узнали, что α - и β -цепи гемоглобина появились в результате генной дупликации. Они произошли от одной-единственной полипептидной цепи, сходной с той, которая была обнаружена в крови миноги. До этого события дупликация гена привела к дивергенции гемоглобина и миоглобина, а после него, также в результате генной дупликации, у предка человека из β -цепи возникли γ - и δ -цепи (δ -цепь входит в состав HbA₂, а γ -цепь является составной частью гемоглобина плода HbF). Что же касается появления новой функции, то последним крупным адаптивным преобразованием такого рода был переход от мономерного гемоглобина к тетрамерному ($\alpha_2\beta_2$), произошедший около 500 млн. лет назад. Благодаря этому переходу стало возможным быстрое отщепление кислорода, связанного с гемоглобином, при понижении рН (в тканях) и быстрое присоединение кислорода с увеличением рН (в легких). Такое функциональное приобретение поставило этих животных в более благоприятные условия (по сравнению с обладателями мономерного гемоглобина) в таких жизненных ситуациях, как бегство от хищника или погоня за жертвой. Известно, что геном человека в норме содержит по крайней мере по одному гену β -, δ - и ζ -цепей гемоглобина и что гены α - и γ -цепей дублицированы (хороший обзор генетики гемоглобинов дается в книге Фогеля и Мотульского [561]). В работе Гуссен-

са и др. [140] приводятся данные о распространенности триплицированных α -глобиновых генов. Согласно этим авторам, у индивидуумов, гетерозиготных по триплицированным α -локусам ($\alpha\alpha/\alpha\alpha\alpha$), не наблюдается каких-либо сопряженных с ними аномальных клинических проявлений. Триплицированные α -глобиновые локусы, вероятно, возникли в результате неравного кроссинговера. У мышей BALB/c полный комплекс генов β -глобина состоит из семи сцепленных локусов [194]. Дубликации генов приводят также к возникновению нефункционирующих «псевдогенов». Согласно недавнему сообщению Нисиоки и др. [421], у мышей имеются по крайней мере два нефункционирующих α -гена: α -3 и α -4. В гене α -3 недостает двух интронов, присутствующих в нормальных функционирующих глобиновых генах, а ген α -4 несет мутации, аналогичные имеющимся в одном из аномальных гемоглобинов человека — гемоглобине М. Нуклеотидная последовательность α -3-гена расшифрована Нисиоки и др. [421]; по сравнению с нормальным α -геном в α -3-гене имеется много вставок и делеций и он не может кодировать молекулу глобина. Недавно была установлена нуклеотидная последовательность α -псевдогена человека (цит. по [474]).

Другое интересное свидетельство того, что гены с новой функцией возникают в результате дубликаций, дает сравнение α -лактальбумина и лизоцима. Лизоцим — фермент, разрушающий клеточную стенку бактерий. Он имеется у самых разных организмов. α -Лактальбумин сам по себе ферментом не является, но входит в состав лактозосинтетазной системы, функционирующей в молочных железах во время лактации. Согласно Дикерсону и Гейсу [81], аминокислотные последовательности лизоцима белка куриного яйца, человека и α -лактальбумина коровы достаточно близки, чтобы происхождение этих белков от общего предка не вызывало никаких сомнений. Кроме того, укладка полипептидных цепей этих белков по существу одинакова, что также подтверждает предположение об их происхождении от общего предкового полипептида. Судя по различиям в аминокислотных последовательностях, лизоцимы белка куриного яйца и человека практически равноудалены от α -лактальбумина коровы ($d_{aa}=79$ и 82), однако расстояние между самими этими лизоцимами ($d_{aa}=57$) намного меньше расстояния от них до α -лактальбумина. С другой стороны, естественно предположить, что дивергенция α -лактальбумина и лизоцима произошла при зарождении линии млекопитающих, когда формировалась система секреции молока. Следовательно, если гипотеза о постоянстве скорости молекулярной эволюции справедлива, то расстояние между α -лактальбумином коровы и лизоцимом человека должно быть намного меньше расстояния между α -лактальбумином коровы и лизоцимом белка куриного яйца, что не соответствует действительности. Для устранения

этого противоречия Дикерсон и Гейс [81] допускают, что гипотеза о постоянстве скорости молекулярной эволюции неверна и что скорость эволюции α -лактальбумина была гораздо выше скорости эволюции лизоцима. Различия в аминокислотных последовательностях этих белков свидетельствуют о том, что для первого из них скорость изменения по крайней мере в два раза больше, чем для второго. По мнению Дикерсона и Гейса, если дублированный ген, изменяясь, приводит к появлению белка с какой-то новой функцией, давление отбора становится чрезвычайно жестким и нуклеотидная последовательность этого гена изменяется очень быстро. Такое объяснение было принято широким кругом исследователей, ссылающихся на него как на установленный факт. Однако выполненная позднее работа Уайта и др. [579] заставила усомниться в его правильности. Эти авторы провели более широкое сравнение, сопоставив, в частности, аминокислотные последовательности α -цепей лактальбуминов человека, морской свинки и коровы, а также лизоцимов человека, крысы и павлиана. Они обнаружили, что скорости эволюции этих двух белков у всех исследованных ими млекопитающих примерно одинаковы, и предложили другое объяснение указанному противоречию, предположив, что дубликация, приведшая к превращению одного из генов лизоцима в ген лактальбумина, произошла раньше, чем дивергенция птиц и млекопитающих. Если это действительно так, то гипотеза о постоянстве скорости молекулярной эволюции остается справедливой.

По моему мнению, для целого ряда случаев гипотеза об очень высоких темпах мутационных замен сразу после дубликации гена остается приемлемой, однако и здесь более естественным представляется объяснение, основанное на предположении о случайной фиксации слабо вредных мутаций, чем на предположении о необычайно высокой скорости адаптивных мутационных замен. Приведенные выше примеры ясно показывают, что дубликация может приводить к появлению нового гена с какой-либо полезной функцией, однако она, как правило, сопровождается также дегенерацией одной из дублированных копий. На самом деле, как отмечали Ней [399], а также Ота и Кимура [453], дубликация, по-видимому, гораздо чаще приводит к возникновению неполноценных, нефункционирующих генов, чем полезных новых генов. Причина этого в том, что многие мутации (включая делеции и вставки нуклеотидов), которые были бы абсолютно вредными до дубликации, после дубликации перестают быть таковыми (становятся нейтральными или лишь очень слабо вредными), а это позволяет им распространяться по популяции путем случайного дрейфа (математическая сторона этого вопроса рассмотрена в работе [248]). Как впервые отметили Мёллер и Бриджес (см. [39]),

важно здесь то, что дубликации позволяют генам «ставить» ранее запрещенные эволюционные эксперименты, освобождая их от давления естественного отбора, который до этого оказывал непреодолимое «стабилизирующее» действие, сохраняя *status quo*. Иными словами, дубликации генов создают условия для гораздо более эффективного влияния случайного дрейфа на динамику мутаций. Это делает возможной фиксацию мутаций, слабо вредных в данных условиях, но могущих оказаться полезными для адаптации к какой-то новой среде. Таким образом, если принять во внимание роль дубликации генов, то окажется, что теория нейтральности имеет непосредственное отношение к проблеме прогрессивной эволюции. Парадоксальное на первый взгляд утверждение Хо и Сондерса [181] о том, что «относительное отсутствие естественного отбора может быть предпосылкой значительного эволюционного прогресса», становится понятным в свете нейтралистской парадигмы.

Гаплоидный набор хромосом млекопитающего содержит примерно в 500 раз больше ДНК, чем геном *E. coli*, что можно считать следствием многократных дубликаций, происходивших в ходе эволюции. То, что дубликации генов должны были играть очень важную роль в прогрессивной эволюции от прокариот до млекопитающих, следует также из тезиса Дэйхоф [77], в соответствии с которым примерно 50 000 белков, имеющих, по оценкам, у человека, можно сгруппировать в 500 надсемейств, каждое из которых включает около 100 белков.

Прежде всего не следует забывать, что в основе всей этой проблематики лежит ключевой вопрос о механизмах фиксации самих дубликаций. Согласно распространенной точке зрения (см., например, [429], с. 65), если некоторый локус представлен парой сверхдоминантных аллелей, скажем A_1 и A_2 , то появление в результате дубликации хромосомы с двумя локусами, содержащими соответственно A_1 и A_2 , может обеспечить несущим ее особям постоянное преимущество, обусловленное их константной гетерозиготностью (без выщепления неприспособленных гомозигот). Отметим, что для рассматриваемой дубликации необходим кроссинговер между гомологичными хромосомами. Спифорд [527] провел специальное популяционно-генетическое исследование, чтобы показать, каким образом такая гетерозисная дубликация может включаться в генофонд вида.

Однако не исключено, что внутривидовая дубликация приводит к образованию хромосомного сегмента типа A_1A_1 . Если такая дубликация селективно нейтральна или только очень слабо вредна, она зафиксирована в популяции в результате случайного генетического дрейфа. (Дальнейшее увеличение числа копий генов, т. е. их трипликация и т. д., может быть результатом неравного кроссинговера между сестринскими хроматидами или неточного синапсиса и кроссинговера го-

мологичных хроматид, как в случае мультигенных семейств, который мы рассмотрим позже.) Подобно обычным точковым мутациям, дубликации генов постоянно, с невысокой частотой возникают в популяциях, и «мутационное давление» внутрихромосомных дубликаций в сочетании со случайным дрейфом в ходе длительной эволюции, по-видимому, было основной причиной широкого распространения повторяющихся структур в геномах высших организмов. До некоторой степени сходное рассмотрение этого вопроса с приведением числовых примеров дано в работе [354].

Тетраплоидизация в отличие от внутрихромосомной дубликации приводит к одновременному удвоению всех локусов. Для растений этот способ образования дублицированных генов довольно обычен, но у животных он встречается очень редко. Тем не менее среди рыб подотряда Salmonidea, к которым относятся, например, форель и лосось, есть тетраплоидные виды, достигшие разных степеней диплоидизации [429]. Аллендорф [2], а также Феррис и Витт [102] провели соответствующие исследования и получили результаты, имеющие самое непосредственное отношение к полемике между сторонниками нейтралистской и селекционистской теорий.

Как у лососевых, так и у чукучановых рыб имеется большое число дублицированных локусов, в которых после тетраплоидизации одна из копий гена утратила свою функцию (вероятно, вследствие фиксации нефункционирующего «нулевого» аллеля, который до дубликации был абсолютно вредным, а после нее перестал быть таковым). На долю таких локусов приходится примерно 50% (эта оценка была получена для дубликаций, происшедших в течение 50 млн. лет). Аллендорф [2] обратил внимание на интересную возможную зависимость между уровнем гетерозиготности и частотой, с которой одна из копий дублицированного гена перестает экспрессироваться. Он указывает, что если гетерозиготность по ферментным локусам поддерживается балансирующим отбором, например вследствие преимущества гетерозигот (сверхдоминирование), то дублицированные локусы обычно будут поддерживаться в состоянии «фиксированной гетерозиготности», как это следует из модели внутрихромосомной дубликации Спофорта [527]. Такая тенденция должна быть выражена в большей степени при более сильном сверхдоминировании и, следовательно, при более высокой средней гетерозиготности. Иными словами, чем выше гетерозиготность, тем меньше вероятность того, что после тетраплоидизации дублицированный ген перестанет экспрессироваться. Однако если аллели, ответственные за полиморфизм ферментов, селективно нейтральны или почти нейтральны и сохраняются в популяции благодаря равновесию между появлением мутаций и их случайной элиминацией в результате дрей-

фа, то большая гетерозиготность будет достигаться за счет более высоких темпов возникновения нейтральных аллелей. Можно предположить, что эффективные скорости мутирования к нулевым и нейтральным аллелям изменяются параллельно. Можно предположить также, что при уменьшении селективных ограничений, налагаемых на молекулу, среди функционирующих мутантных аллелей увеличивается доля нейтральных (не вредных) аллелей и в то же время, вероятно, уменьшается неблагоприятный эффект нефункционирующих нулевых мутантных аллелей, что приводит к аналогичным последствиям. В этом случае утрата одной из копий дублированных генов способности к экспрессии после тетраплоидизации должна происходить тем чаще, чем выше уровень гетерозиготности.

Имея в виду два указанных альтернативных предположения, Аллендорф [2] проанализировал соответствующие данные по белкам лососевых и чукучановых. Оказалось, что для 12 белков радужной форели средняя гетерозиготность ($\bar{H}$) по локусам, сохранившим способность обеих копий дублированного гена к экспрессии, составляет 6%, а по локусам, утратившим ее, — 15%. Иными словами, чем выше гетерозиготность, тем скорее дублированный ген утрачивает способность к экспрессии, что находится в соответствии с теорией нейтральности.

Сходные результаты были получены Феррисом и Виттом [102] на основе данных о 12 ферментах 26 видов чукучановых рыб. По мнению Аллендорфа, эти результаты также подтверждают теорию нейтральности. Такахата и Маруяма [543] провели математический анализ процесса, приводящего в ходе эволюции к фиксации нулевых аллелей в независимых дублированных локусах. Получив численными методами решения двумерного обратного уравнения Колмогорова, они оценили скорость потери одного из функционирующих генов и величину гетерозиготности в заданный момент времени при разных типах отбора, включая случаи, когда нулевые аллели полностью рецессивны (т. е. вредны только в двойной гомозиготе), частично рецессивны и сверхдоминантны. Оказалось, что при полной рецессивности время, необходимое для фиксации нулевого аллеля в одном из дублированных локусов с вероятностью 50%, должно быть намного меньше 50 млн. лет. Кроме того, родственные виды должны быть высоко гетерозиготными по тем локусам, в которых потеря функции еще не завершилась. Все это не соответствует фактам. В случае сверхдоминирования фиксация замедляется, но уровень гетерозиготности остается очень высоким, что опять же противоречит имеющимся данным. Единственное реалистичное допущение, согласующееся с фактами, — предположение об очень слабой вредности нулевых аллелей во всех гетерозиготных комбинациях, для которых $N_e s$

(произведение эффективного размера популяции на коэффициент отбора против гетерозигот, содержащих по крайней мере один нулевой аллель) равно примерно единице. Важно здесь то, что после дупликации нулевые аллели, несмотря на их очень слабые вредные эффекты, могут фиксироваться в результате случайного дрейфа.

Я не могу завершить это обсуждение эволюционной роли дупликации, не сказав о недавно сформулированной концепции «мультигенного семейства» [184]. Такое семейство представляет собой группу тандемно повторяющихся генов, характеризующихся четырьмя признаками: множественностью, тесным сцеплением, гомологией последовательностей и родственными или частично дублирующими друг друга фенотипическими функциями. В мультигенные семейства объединяются, например, гены иммуноглобулинов, гены рибосомных РНК (рРНК) и гены гистонов. Эти гены многократно повторяются в геноме и в эволюционном плане характеризуются двумя интересными особенностями, а именно: быстрым изменением размера семейства и согласованной эволюцией. Последний из упомянутых феноменов называют также горизонтальной эволюцией. Я поясню суть этого крайне интересного, но загадочного явления на примере спейсерных участков, ассоциированных с генами рРНК эукариот (подробности см. в работе [184]). У африканской шпорцевой лягушки *Xenopus laevis* ген, кодирующий 18S- и 28S-рРНК, повторяется в отдельном участке хромосомы (ядрышковом организаторе) примерно 450 раз, а каждая повторяющаяся единица состоит из гена, кодирующего 18S- и 28S-РНК, а также спейсерных участков. Практически такое же число повторов, имеющих такую же структуру, выявлено у близкородственного вида, *Xenopus mulleri*. Однако при сравнении нуклеотидных последовательностей у этих двух видов обнаруживается, что их 18S- и 28S-гены идентичны, а спейсерные участки значительно дивергировали. Так как спейсеры не выполняют никаких явных функций, то с точки зрения теории нейтральности логично предположить, что они подвергаются менее жестким селективным ограничениям и поэтому быстрее накопили мутации вследствие случайного дрейфа. Однако практически полная идентичность нуклеотидных последовательностей спейсерных участков в пределах каждого вида представляется загадочной. Как же поддерживается такая внутривидовая гомогенность спейсерных участков при столь высоких темпах их эволюции?

Остроумное объяснение этого феномена, основанное на предположении о гомологичном неравном кроссинговере, предложили Смит [519], а также Блэк и Гибсон [22] (см. также [544]). Рис. 5.1 иллюстрирует процесс, приводящий к гомогенности некоего мультигенного семейства, локализованного на

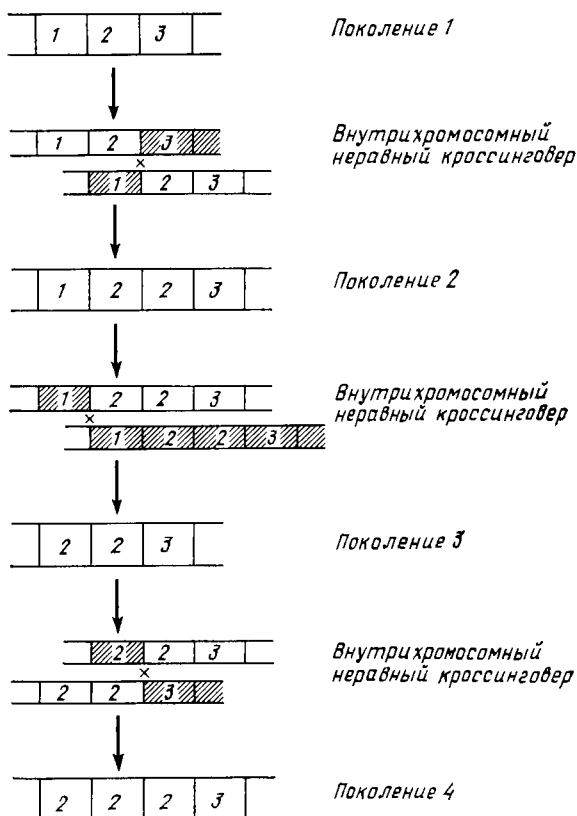


Рис. 5.1. Диаграмма, иллюстрирующая процесс согласованной эволюции, происходящей с помощью неравного кроссинговера, который приводит к случайной фиксации гена 2 в соответствующем участке хромосомы.

одной хромосоме, с помощью повторяющегося «внутрихромосомного» (т. е. происходящего между сестринскими хроматидами) неравного кроссинговера. Как отмечали Худ и др. [184], а также Смит, этот процесс аналогичен случайной фиксации одного из аллелей в популяции конечной численности, и поэтому было предложено называть его «кроссоверной фиксацией». Среднее время, необходимое для фиксации одной из вариантных копий семейства в «родословной» хромосомы, называют временем кроссоверной фиксации. Смит, а также Блэк и Гибсон провели имитационные эксперименты, с тем чтобы количественно оценить число кроссоверов, необходимое для достижения определенного уровня гомогенности членов семейства. Следует отметить, что благодаря внутрихромосомному неравному крос-

синговеру мутация, возникшая в одном из членов мультигенного семейства, может распространиться случайным образом по всему семейству (горизонтальное распространение мутации). Поэтому соседние члены семейства оказываются генетически более сходными, чем локализованные в более удаленных друг от друга участках хромосомы (результаты соответствующего математического анализа приведены в работе [265]). Одновременно с генетическим дрейфом, благодаря которому становится возможной случайная фиксация одной из копий гомологичных мультигенных семейств, протекает согласованная эволюция. Соответствующая математическая теория, основанная на принципах популяционной генетики, получила интенсивное развитие в работах Оты [439, 443—446]. Выполненные ею разнообразные исследования, посвященные эволюции и изменчивости мультигенных семейств, недавно были представлены и систематизированы в превосходной монографии [447]. В этой книге процесс эволюции мультигенного семейства рассматривается как процесс двойной диффузии: диффузии мутаций по хромосоме в результате неравного кроссинговера и диффузии таких хромосом в популяции вследствие случайного генетического дрейфа. Были рассмотрены также внутри- и межхромосомные коэффициенты идентичности (вероятности идентичности генов). Особый интерес представляет приложение теории Оты к анализу проблемы разнообразия аминокислотных последовательностей иммуноглобулинов — очень хорошо изученного мультигенного семейства. Она обнаружила, что каркасные (framework) участки, расположенные внутри вариабельных областей иммуноглобулинов, эволюционировали (накапливали аминокислотные замены) немного быстрее гемоглобинов, тогда как гипервариабельные участки, входящие в состав сайтов, связывающих антигены, эволюционировали приблизительно с такой же скоростью, что и фибринопептиды. Кроме того, высокая внутривидовая изменчивость гипервариабельных участков оказалась примерно на том уровне, который был получен теоретически с учетом высоких темпов их эволюции. Интересно, что метод диффузионных уравнений и нейтралистская парадигма оказались полезными для выяснения и количественного описания природы мультигенных семейств.

Типичное мультигенное семейство состоит из множества повторяющихся структур (так, например, ген 5S-рРНК повторяется около 2400 раз у *X. laevis* и примерно 9000 раз у *X. mul-leri*), но с точки зрения эволюционной и популяционной генетики большой интерес представляют также небольшие мультигенные семейства. Циммер и др. [618] сообщили, что у пяти видов человекообразных обезьян (шимпанзе, карликового шимпанзе, гориллы, orangутана и гиббона) гены α -цепи гемоглобина взрослых особей, как и у человека, дублированы и тесно

сцеплены и что в той области, где локализованы эти гены, повидимому, осуществлялись быстро протекавшие циклы дупликации и потери генов. То, что в данном случае протекала эффективная согласованная эволюция, с очевидностью следует из того факта, что в пределах видов α -цепи почти идентичны. По оценкам Циммера и др., внутри этих видов сходство α -полипептидов примерно в 10 раз выше, чем должно было бы быть, если бы имела место дивергентная эволюция в течение 10 млн. лет (по мнению авторов, именно столько времени прошло с момента появления общих предков этих человекообразных обезьян). Кроме того, основываясь на результатах собственного исследования с использованием пяти рестриктаз, они обнаружили у шимпанзе полиморфизм по числу α -генов, причем у пяти изучавшихся особей шимпанзе преобладающим оказалось трехгенное, а не двухгенное состояние. Иная картина наблюдалась при изучении соответствующей системы полиморфизма у человека: согласно Гуссенсу и др. [140], индивидуальные α -локусы, находясь в меньшинстве; большинство же имеет дуплицированные α -гены. Математическое обоснование теории согласованной эволюции в небольших мультигенных семействах было разработано Отой [448] с использованием двух моделей. Применив свою теорию к упомянутым здесь данным, она нашла, что частота неравного кроссинговера, вероятно, равна примерно 10^{-4} на поколение, а коэффициент отбора против хромосом с тремя α -генами или одним α -геном составляет примерно 10^{-2} . Она также отметила, что нефункционирующие гены могут возникать в результате неравного кроссинговера и, если такие гены селективно нейтральны, они будут накапливаться в популяции вследствие случайного дрейфа.

Позднее было выдвинуто предположение, что кроме неравного кроссинговера важную роль в протекании согласованной эволюции играет конверсия генов [17]. Конверсия гена, грубо говоря, приводит к дупликации одной генной копии и одновременно вызывает потерю другой. Таким образом, она эквивалентна одному циклу неравного кроссинговера с дупликацией и делецией [21, 441]. Основное отличие результата конверсии от последствий неравного кроссинговера заключается в том, что она не сопровождается уменьшением или увеличением мультигенного семейства, т. е. не приводит к изменению числа генов в семействе.

При наличии направленности (или «полярности») генной конверсии процесс межвидовой дивергенции мультигенных семейств значительно ускоряется. Доувер [90] назвал этот процесс и процесс случайной конверсии «молекулярным драйвом» и подчеркнул, что данное явление существенно для эволюции некоторых кодирующих и некодирующих семейств нуклеотид-

ных последовательностей у дрозофилы, рассеянных по нескольким хромосомам. Математический анализ распространения мутантных генов посредством направленной генной конверсии в пределах отдельного генома был проведен Нагилаки и Петесом [396], однако теория молекулярного драйва на уровне популяции пока не разработана.

5. Селективная элиминация вредных мутаций и случайная фиксация селективно нейтральных или очень слабо вредных мутаций происходят в ходе эволюции гораздо чаще, чем положительный дарвиновский отбор благоприятных мутаций.

Это положение представляет собой расширенную формулировку исходной теории нейтральности и основано на тезисе, выдвинутом Отой [435, 437]. Согласно этому тезису, для молекулярной эволюции важны не только селективно нейтральные, но и слабо вредные мутации. Адаптивные изменения, обусловленные положительным дарвиновским отбором, на молекулярном уровне (не говоря уже о фенотипическом уровне) несомненно происходят, однако я считаю, что абсолютно благоприятные мутационные замены составляют — по сравнению с относительно большим числом мутационных замен «недарвиновского» типа — меньшинство, т. е. фиксация мутантных аллелей в популяции осуществляется в результате случайного дрейфа генных частот. Отсюда следует важное для теории нейтральности положение о том, что *«нейтральные мутации» не являются крайними формами селективно благоприятных мутаций, однако их можно рассматривать как крайние формы вредных мутаций, когда их влияние на приспособленность становится бесконечно малым.* Это означает, что мутации вызывают эволюционные изменения всякий раз, когда ослабевает действие отрицательного отбора.

Чтобы проиллюстрировать это положение, рассмотрим случай высокой скорости эволюции инсулина морской свинки. Для инсулина (А- и В-цепи), вообще говоря, характерна довольно низкая скорость эволюции: $(0,3 \sim 0,4) \cdot 10^{-9}$ на аминокислотный сайт в год. Однако у морской свинки, а также у ее близкого «родственника» нутрии (оба животных принадлежат к группе «гистрикоморфных грызунов») этот белок эволюционировал очень быстро: скорость замен почти в десять раз превышала приведенную выше величину (см. [268]; более подробные данные см. в работе [203]). С точки зрения неodarвинизма естественно считать такое быстрое эволюционное преобразование результатом адаптивного изменения под действием естественного отбора. Даже Кинг и Джукс [268] в своей статье «Недарвиновская эволюция» объясняли этот случай быстрого эволюционного изменения, привлекая представление о «положительном естественном отборе». Однако, как отмечали Кимура и Ота

[261], гораздо более вероятно, что в процессе видообразования инсулин морской свинки просто освободился от налагаемых на него селективных ограничений. Это сделало возможным накопление мутаций, которые раньше были бы забракованы. Такой вывод получил подтверждение в работе Бландела и др. [23], изучавших трехмерную структуру молекул инсулина. Согласно этим авторам, в клетках морской свинки, вырабатывающих инсулин, отсутствует цинк (что сопровождается потерей обычно инвариантного гистидина В10). Это указывает на возможность радикального изменения третичной структуры данного белка. Логично предположить, что снятие ограничения, налагаемого взаимодействием инсулина морской свинки с цинком, привело к очень быстрому накоплению мутаций в этой молекуле. Наше предположение согласуется также с тем фактом (соответствующие данные приведены, например, в работе [203]), что скорость аминокислотных замен для инсулинов морской свинки и нутрии примерно такая же, как для С-пептида проинсулина, который во всех случаях эволюционировал с очень высокой скоростью и для которого, как считается, характерны гораздо меньшие функциональные ограничения, чем для инсулина. Кроме того, нейтралистская интерпретация, по-видимому, подтверждается данными об очень низкой биологической активности инсулина морской свинки [24], составляющей около 2% от активности инсулина крупного рогатого скота. Согласно недавнему сообщению Хорука и др. [186], очень низкая биологическая активность характерна для инсулина всех гистрикоморфных грызунов, однако Бландел и Вуд [24] придерживаются мнения, что инсулин морской свинки возник в результате приспособления к локальному дефициту цинка, в условиях которого жили ее предки.

Другой интересный случай, который можно объяснить случайной фиксацией слабо вредных, почти нейтральных мутаций, — утрата человеком и некоторыми другими видами способности синтезировать аскорбиновую кислоту. Джукс и Кинг [209] предположили, что это произошло в результате какого-то «нейтрального эволюционного изменения». Согласно Джуксу и Кингу, способность синтезировать аскорбиновую кислоту (витамин С) — характерный признак наземных позвоночных. Однако он отсутствует не только у человека, но и у обезьян, морских свинок, питающихся фруктами рукокрылых и у некоторых воробьиных птиц. Эти животные потребляют пищу, богатую аскорбиновой кислотой, и поэтому можно предположить, что мутации, приводившие к потере способности синтезировать аскорбиновую кислоту, не были особенно вредными и при постоянном мутационном давлении фиксировались в результате случайного дрейфа генных частот. Согласно другой гипотезе, потеря способности синтезировать аскорбиновую кислоту носила приспособ-

собительный характер и произошла под действием положительного естественного отбора. Однако гипотеза Джукса и Кинга представляется более правдоподобной, так как многие виды травоядных позвоночных, потребляющих пищу с высоким содержанием аскорбиновой кислоты, тем не менее могут ее синтезировать. Их гипотезу подтверждает также тот факт, что способность синтезировать аскорбиновую кислоту была утрачена видами из разных ветвей филогенетического дерева. Математическая трактовка вопросов, связанных с оценкой среднего времени фиксации нейтрального или слабо вредного аллеля в популяции конечной численности вследствие случайного дрейфа при постоянном давлении мутаций, дана в работе Кимуры [240].

Рассматривая действие естественного отбора на молекулярном уровне, следует иметь в виду, что селективные ограничения (обычно в результате действия отрицательного отбора, т. е. элиминации функционально вредных изменений) налагаются не столько на первичную структуру молекул (аминокислотную последовательность), сколько на их вторичную, третичную и четвертичную структуры. Наличие селективных ограничений, часто устанавливаемое по неслучайности распределения аминокислот или оснований, не противоречит теории нейтральности. То, что случайный дрейф может происходить с сохранением фенотипического однообразия, с точки зрения здравого смысла, возможно, и не очевидно, однако этот скрытый процесс поддается математическому анализу (см. разд. 6.8).

Если эволюционное изменение обусловлено главным образом мутационным давлением, верхний предел скорости эволюции определяется общей скоростью мутирования. В этой связи интересно отметить, что высокая скорость эволюции инсулина морской свинки не так уж сильно отличается от скорости эволюции фибринопептидов, а последняя в свою очередь примерно равна скорости нуклеотидных замен в ДНК млекопитающих, большая часть генома которых не содержит функционирующих генов (см. [453]). Следует также отметить, что накопление очень слабо вредных мутаций в ходе случайного дрейфа по существу приводит к таким же последствиям, что и ухудшение окружающей среды, и поэтому время от времени для спасения вида от вымирания должны происходить адаптивные генные замены. Хочу подчеркнуть еще раз: приведенные здесь доводы вовсе не означают, что на уровне информационных макромолекул адаптивные изменения не происходят. Напротив, удивительные функции молекулярных механизмов, от которых зависит сама жизнь, должны быть продуктом положительного дарвиновского отбора. Однако большинство изучаемых в настоящее время молекул, таких, как гемоглобины, цитохром *c*, транспортные РНК и т. д., должны были приобрести все существен-

ные особенности своей структуры еще в далеком прошлом; последующие изменения представляли собой главным образом «вариации на одну и ту же тему». Это ясно видно, например, из того факта, что все молекулы глобинов, включая гемоглобины насекомых и морских кольчатых червей, имеют совершенно одинаковую третичную структуру с характерной «миоглобиновой укладкой» (см. [81]), хотя их первичная структура существенно различается. Время от времени появляются пространные публикации, содержащие «адапционистское» истолкование тех или иных аминокислотных замен (см., например, работу [138], посвященную эволюции глобинов), однако к таким интерпретациям необходимо относиться с осторожностью.

В общем случае адаптивные мутационные замены должны представлять собой ответ организма на «вызов», брошенный окружающей средой. Таким образом, безусловно благоприятных мутаций не существует; «благоприятные мутации» образуют весьма специфический, небольшой класс мутаций, возникших из очень большой совокупности вредных мутаций. Как это ни парадоксально звучит, можно, по-видимому, сказать, что благоприятные мутации представляют особую разновидность вредных мутаций. Способы образования благоприятных мутаций столь разнообразны, что найти какие-либо простые правила, их описывающие, очень нелегко. В данном случае может оказаться полезной одна из теорий Оты [434]. Согласно этой теории, мутации, благоприятные при всех условиях, задаваемых окружающей средой, практически невозможны. С другой стороны, если среда однородна, мутация имеет большие шансы стать благоприятной. Таким образом, вероятность того, что некая мутация будет благоприятной для вида в целом, увеличивается в более однородной окружающей среде. Так как популяции меньшей численности, по-видимому, расселяются по менее гетерогенным средам, то из этой теории следует, что быстрая эволюция на фенотипическом уровне происходит у видов с популяциями малых размеров.

Хотя сколько-нибудь подробное рассмотрение процесса видообразования выходит за рамки этой книги, следует отметить, что проблема случайной фиксации нейтральных или слабо вредных аллелей находится в очевидной связи с проблемой видообразования. Ряд авторов (см., например, [86; 401]) указывали, что для развития репродуктивной изоляции необходимо возникновение мутационных изменений, как правило, в двух или большем числе локусов. Говоря кратко, в становлении репродуктивной изоляции должны участвовать мутантные аллели разных локусов, безвредные (точнее, нейтральные) по отдельности, но вредные в сочетании друг с другом.

Рассмотрим одну простую модель видообразования, предложенную Добржанским [83, 86]. Пусть A и A' — пара аллелей

первого локуса, а B и B' — пара аллелей второго локуса. Предположим, что популяция, имеющая исходный генетический состав $AABB$, разбивается на две географически изолированные субпопуляции; в одной из них в результате случайного дрейфа фиксируется мутантный аллель A' (в предположении селективной эквивалентности A и A' в отсутствие B'), что приводит к образованию локальной расы $A'A'BB$, тогда как в другой субпопуляции фиксируется аллель B' и появляется раса $AAB'B'$. В таком случае скрещивание $A'A'BB \times AAB'B'$, если оно происходит, приводит к появлению особей $A'AB'B$ с пониженной жизнеспособностью, обусловленной эпистатическим взаимодействием между генами A' и B' , что означает становление репродуктивной изоляции. Позднее, при устранении географического барьера между этими двумя популяциями, они могут эволюционировать как симпатрические виды.

Естественный отбор: определение, типы, механизмы действия

6.1. Значение естественного отбора

Для правильной оценки теории нейтральности очень важно иметь верное представление о той роли, которую играет естественный отбор в изменении и поддержании частот генов в природных популяциях. Поэтому я кратко остановлюсь на определении естественного отбора, типах и механизмах его действия.

Во введении к своей книге «Происхождение видов» Дарвин [72] писал: «Так как рождается гораздо более особей каждого вида, чем сколько их может выжить, и так как, следовательно, постоянно возникает борьба за существование, то из этого вытекает, что всякое существо, которое в сложных и нередко меняющихся условиях его жизни хотя незначительно изменится в направлении для него выгодном, будет иметь более шансов выжить и таким образом подвергнется *естественному отбору*. В силу строгого принципа наследственности отобранная разновидность будет стремиться размножаться в своей новой и измененной форме».

Хотя со времен Дарвина биология достигла очень больших успехов, теория естественного отбора все еще остается единственной приемлемой с научной точки зрения теорией, объясняющей, почему организмы так хорошо приспособлены к среде обитания. Популяционную генетику можно рассматривать в значительной мере как попытку описать на количественном уровне действие естественного отбора и его последствия (опираясь на генетику и используя методы биометрии).

С точки зрения генетики естественный отбор можно определить как дифференциальное размножение вариантов генов [391]. Другими словами, мы можем говорить, что естественный отбор вступает в силу в тех случаях, когда генотипы различаются по приспособленности. В этом смысле термин «приспособленность» отражает способность организма оставлять потомство. Для каждого генотипа приспособленность измеряется числом потомков на особь. Вообще говоря, значительная часть молоди может погибнуть, не достигнув половозрелости (например, у насекомых или рыб), поэтому уместнее говорить о числе

потомков, которые достигают этого возраста. Важно, чтобы число потомков подсчитывалось на той же стадии развития, что и для родительской популяции. В популяциях человека приспособленность удобно измерять числом дочерей на мать. Подробнее с вопросом о природе и мере приспособленности можно ознакомиться в работе Кроу и Кимуры [67].

Однако естественный отбор действует на фенотипы (через выживаемость и воспроизводство), а не на генотипы. Более того, механизмы действия естественного отбора в конкретных ситуациях настолько разнообразны и неоднозначны, что часто их детали остаются для нас совершенно неясными. Чтобы преодолеть эту трудность и выработать единый взгляд на этот вопрос, были предложены несколько фундаментальных концепций и обширная терминология.

Простой и полезный способ классификации естественного отбора основан на выделении двух его типов — положительного и отрицательного. Мы говорим, что действует положительный естественный отбор, если возникшая в популяции мутантная форма гена (называемая просто мутантом) увеличивает среднюю приспособленность обладающих ею особей, способствуя таким образом распространению в популяции. Именно этот тип естественного отбора Дарвин считал главной движущей силой адаптивной эволюции в биологическом мире; такой отбор с полным правом можно назвать дарвиновским. Необязательно, чтобы отобранный ген представлял собой новую мутантную форму; это может быть и такой ген, который вначале обладал неблагоприятным эффектом, но при изменении среды оказался благоприятным.

Вместе с тем если в результате мутации уменьшается приспособленность особи из-за уменьшения ее жизнеспособности и плодовитости, то отбор стремится элиминировать ее. Такой тип отбора может быть назван отрицательным. Говоря о нем, Левонтин [302] применил термин «очищающий».

Несмотря на большое биологическое значение положительного отбора, в природе он наблюдается редко. К немногочисленным хорошо известным и постоянно цитируемым примерам относятся индустриальный меланизм у бабочек [217—219], устойчивость к ДДТ у насекомых и к антибиотикам у бактерий. Примеров отрицательного отбора, напротив, довольно много; неблагоприятные эффекты и пороки развития принято связывать с мутациями, поэтому, как правило, многие учебники изобилуют примерами именно такого рода. Кроме того, интенсивное исследование рецессивных леталей и вредных мутаций в популяциях *Drosophila* однозначно показало, что большинство этих мутантных генов безусловно неблагоприятно как в гомозиготном, так и в гетерозиготном состояниях [376, 380].

Естественный отбор действует в первую очередь на феноти-

пы и лишь опосредованным образом — на генотипы. Связь между фенотипом и генотипом часто является косвенной, особенно если фенотипический признак определяется многими генами, как в случае количественных признаков (таких, как рост человека или вес мышей), на которые большое влияние оказывают и условия среды. Поэтому удобно рассматривать отбор на двух уровнях — фенотипическом и генотипическом.

6.2. Фенотипический отбор по количественным признакам

Рассмотрим вначале фенотипический отбор. Следуя Мазеру [343], отбор по количественным признакам подразделяют на три типа: *стабилизирующий*, *направленный* и *дизруптивный*. Предположим, что для какого-то количественного признака фенотипа (X) характерно некое колоколообразное распределение $F(X)$ (например, нормальное распределение). Отбор на-

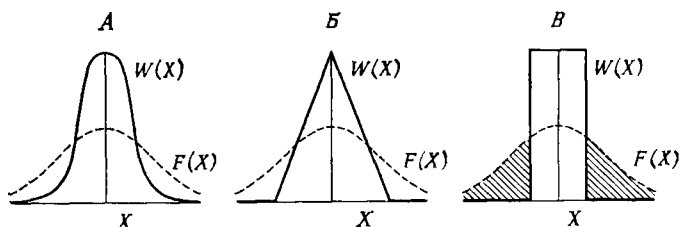


Рис. 6.1. Три примера функций приспособленности при стабилизирующем отборе. Наиболее приспособленными являются особи, у которых значение фенотипического признака X близко к среднему; особи, далекие от оптимума, отсекаются отбором. Функции приспособленности $W(X)$ изображены сплошными кривыми, а функции частоты $F(X)$ — штриховыми. Особи, попадающие на рис. В в заштрихованные области, полностью элиминируются.

зывается *стабилизирующим*, если особи с величиной признака, близкой к средней, имеют наивысшую приспособленность (оптимум), а тех особей, которые по этому признаку отклоняются в положительном или отрицательном направлении от средней, отбор элиминирует. Другими словами, такой отбор отсекает крайние варианты. Несколько примеров стабилизирующего отбора проиллюстрировано на рис. 6.1, где функция приспособленности $W(X)$ изображена сплошной кривой, а функция частоты $F(X)$ — штриховой. Значение стабилизирующего отбора иллюстрирует тот факт, что он является наиболее распространенным типом естественного отбора. Одним из самых известных его примеров может служить связь между весом новорожденных и их смертностью у человека, исследованная Карн и

Пенроузом [212]. Они показали, что у новорожденных, вес которых близок к среднему, выживаемость самая высокая, а у тех, кто отклоняется от среднего в любом направлении, наблюдается тенденция к повышенной смертности. Более подробный анализ естественного отбора в популяциях человека читатель может найти у Кавалли-Сфорца и Бодмера [41]. Другие примеры стабилизирующего отбора приведены, в частности, в работах Мазера и Уоллеса [564]. Описывая этот тип отбора, Холдейн [165] предпочел термин *центростремительный отбор*, введенный ранее Симпсоном [512], так как термин «стабилизиру-

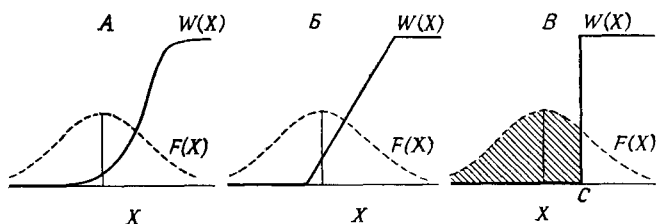


Рис. 6.2. Три примера функций приспособленности при направленном отборе. Сплошные кривые — функции приспособленности, штриховые — функции частоты. Особи, представленные на рис. В заштрихованными участками, полностью элиминируются.

ющий» иногда используют для обозначения другого типа отбора, который удерживает развитие особи в нормальном русле в неблагоприятных условиях внешней среды. Этот последний тип отбора Уоддингтон [562] назвал также *канализирующим отбором*. С другой стороны, в том же смысле, что и центростремительный отбор, нередко используют термин *нормализующий отбор*. Стабилизирующий отбор характеризуется тем, что в результате его действия дисперсия признака уменьшается, а среднее значение практически не изменяется.

Рассмотрим теперь *направленный отбор*. Если оптимальное значение фенотипического признака существенно отличается от среднего, то отбор стремится сдвинуть эту величину в направлении к оптимуму. Этот тип отбора был назван направленным, а Симпсон [512] и Спасс [524] употребили также термин «линейный отбор». Дарвин предположил существование этого типа отбора, экстраполируя результаты искусственного отбора для объяснения эволюционных изменений в природе. Теперь мы знаем, что направленный отбор встречается гораздо реже, чем стабилизирующий. Его действие должно в основном ограничиваться такими ситуациями, когда вид оказывается перед каким-то новым «вызовом» со стороны внешней среды. Самым подходящим примером служит индустриальный меланизм у

Lepidoptera. Полезную информацию для анализа направленного отбора дает также рассмотрение искусственного отбора, например при селекции животных на высокую удойность. На рис. 6.2 приведены три примера функции приспособленности (сплошные кривые) при направленном отборе. Рис. В иллюстрирует «отсекающий отбор», который обычно используется при селекции животных. Точка С на оси абсцисс — это граничная точка; при размножении используют только тех особей, у которых значение фенотипического признака больше С. Заштри-

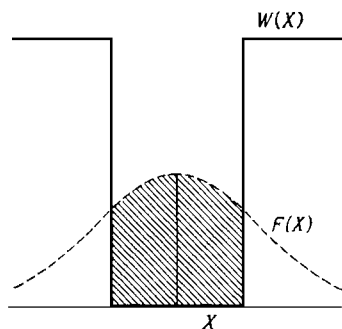


Рис. 6.3. Пример дизруптивного (центробежного) отбора. Наиболее приспособленными являются особи, которым соответствуют концы распределения, а особи с промежуточным значением фенотипического признака (заштрихованная площадь) полностью элиминируются.

хованная область на этом рисунке соответствует доле тех особей, которым не удастся оставить потомство. Однако в природе такой резкий переход едва ли встречается. Более вероятны ситуации, представленные на рис. А и Б.

Как известно, отсекающий отбор является наиболее эффективной формой направленного отбора. Удивительно, что, когда отбор модифицируется так, что приспособленность более или менее линейно возрастает, превышая одно или два стандартных отклонения величины селектируемого признака, эффективность отбора хотя и уменьшается, но весьма незначительно [69].

Третий тип фенотипического отбора — дизруптивный отбор — наблюдается в тех случаях, когда в одной популяции существуют два (или более) оптимума. Пример отбора этого типа приведен на рис. 6.3; здесь отбираются оба края распределения, а все, что находится посередине, элиминируется. Такая ситуация возможна, если в среде обитания вида имеются две ниши и каждый из краев распределения «приспособлен» к своей нише. Этот случай противоположен центростремительному отбору и, следовательно, может быть назван центробежным отбором. Согласно Холдейну [165], центробежный отбор наблюдается очень редко, за исключением случаев межвидовых скрещиваний.

6.3. Генотипический отбор

В прямом количественном анализе действия естественного отбора на генетические изменения в природных популяциях не рассматривается фенотипический отбор, а определяются значения приспособленности непосредственно для каждого генотипа и затем рассчитываются изменения генных частот.

Проанализируем для примера динамику частот двух аллелей A_1 и A_2 в очень большой гаплоидной популяции с дискретными неперекрывающимися поколениями. Пусть W_1 и W_2 — приспособленности A_1 и A_2 , p — частота A_1 , а q — частота A_2 . В следующем поколении частота A_1 будет равна

$$p' = pW_1/\bar{W}, \quad (6.1)$$

где $\bar{W} = pW_1 + qW_2$ — средняя приспособленность популяции. Изменение частоты при переходе от одного поколения к следующему ($p' - p$), обычно обозначаемое через Δp , равно

$$\Delta p = (W_1 - W_2)pq/\bar{W}. \quad (6.2)$$

Часто бывает удобнее выражать Δp через коэффициенты отбора. Пусть $1+s$ и 1 — относительные приспособленности A_1 и A_2 , так что $1+s = W_1/W_2$; тогда

$$\Delta p = sp(1-p)/\bar{w}, \quad (6.3)$$

где $\bar{w} = 1+sp$, а s — коэффициент отбора для A_1 . Если s по абсолютной величине мало, знаменатель $\bar{w}$ будет очень близок к единице и в уравнении (6.3) им можно пренебречь. Кроме того, можно показать [237], что если принять во внимание регуляцию численности популяции, то приближенная формула $\Delta p = sp(1-p)$ будет справедлива в более широком диапазоне условий, чем утверждалось выше. Вообще говоря, если отбор не очень интенсивен, процесс направленного изменения частоты гена в гаплоидной популяции можно с достаточной степенью точности описать дифференциальным уравнением

$$\frac{dp}{dt} = sp(1-p), \quad (6.4)$$

где t — время. В этом уравнении t — непрерывная переменная, но за единицу ее измерения можно принять одно поколение.

Часто удобно провести логит-преобразование частоты гена p , т. е.

$$z = \ln \frac{p}{1-p}. \quad (6.5)$$

Новая переменная z — это натуральный логарифм отношения частот генов. Она непрерывно меняется от $-\infty$ до $+\infty$ при

изменении p от 0 до 1. Тогда уравнение (6.4) сводится к

$$\frac{dz}{dt} = s(t), \quad (6.6)$$

где $s(t)$ — значение s в момент времени t . Это уравнение легко решить, представив z как функцию времени t :

$$z_t = z_0 + \int_0^t s(\tau) d\tau. \quad (6.7)$$

Заметим, что если $s(\tau) = s$, т. е. является константой, то $z_t - z_0 = st$, или $t = (z_t - z_0)/s$. При этом частоту аллеля A_1 можно получить в явном виде:

$$p_t = e^{zt} / (e^{zt} + 1). \quad (6.8)$$

В качестве примера рассмотрим ситуацию, когда A_1 имеет постоянное 1%-ное селективное преимущество над A_2 , так что $s(t) = 0,01$, и зададим вопрос: сколько потребуется времени для того, чтобы частота гена изменилась от 0,5 до 99,5%? Поскольку это соответствует изменению z от $-5,293$ до $+5,293$, время равно $10,586/0,01$, или $1058,6$ поколениям. Такое одностороннее изменение частоты гена также можно назвать «направленным отбором», применив ту же терминологию, что и при описании аналогичной ситуации с фенотипическим отбором.

Уравнение (6.6) можно использовать и для описания процесса «генного отбора» в диплоидной популяции со случайным скрещиванием. Это такой тип отбора, при котором нет доминирования по приспособленности, т. е. приспособленность гетерозиготы занимает промежуточное положение между приспособленностями двух гомозигот. В этом случае относительные приспособленности A_1A_1 , A_1A_2 и A_2A_2 можно обозначить через $1+s$, 1 и $1-s$. Тогда, допустив, что $|s|$ мало, получим с хорошей степенью точности то же самое уравнение (6.6).

Однако на самом деле в диплоидной популяции ситуация гораздо сложнее даже при случайном скрещивании. Поэтому я кратко суммирую несколько случаев, важных для понимания материала последующих частей этой книги. Более подробное изложение с разнообразными примерами читатели могут найти у Кроу и Кимуры [63] и Райта [601].

Рассмотрим диплоидную популяцию со случайным скрещиванием, в которой сегрегирует пара аллелей A_1 и A_2 . Обозначим относительные приспособленности генотипов A_1A_1 , A_1A_2 и A_2A_2 через w_1 , 1 и w_2 . (Будем использовать строчную букву w для обозначения относительной приспособленности, а прописную букву W оставим для абсолютной приспособленности.) Если A_1 имеет несомненные преимущества перед A_2 , так что $w_1 \gg$

≥ 1 и $\omega_2 \leq 1$ (исключая нейтральную ситуацию, когда $\omega_1 = \omega_2 = 1$), то его частота увеличивается. В отсутствие обратных мутаций аллель A_2 в конце концов будет утрачен популяцией. Сходным образом, если преимущество имеет A_2 , будет увеличиваться его частота. Два этих случая представляют собой направленный отбор на генотипическом уровне.

Рассмотрим также несколько примеров, в которых Δp меняет знак в интервале частот $(0,1)$, что означает наличие положения равновесия. Один из наиболее известных случаев — это сверхдоминирование, когда гетерозигота имеет большую приспособленность, чем любая из гомозигот ($0 \leq \omega_1 < 1$ и $0 \leq \omega_2 < 1$). Пусть $\omega_1 = 1 - s_1$ и $\omega_2 = 1 - s_2$; тогда s_1 — это коэффициент отбора против A_1A_1 , а s_2 — коэффициент отбора против A_2A_2 ($s_1 > 0$, $s_2 > 0$). Формула для изменения частоты A_1 за одно поколение имеет вид

$$\Delta p = \{s_2 - (s_1 + s_2)p\}(1-p)/\bar{w}, \quad (6.9)$$

где

$$\bar{w} = 1 - s_1p^2 - s_2(1-p)^2. \quad (6.10)$$

Равновесную величину p , которую мы будем обозначать через $\hat{p}$ и которая удовлетворяет условию $\Delta p = 0$, можно получить, положив, что $s_2 - (s_1 + s_2)p = 0$. Отсюда

$$\hat{p} = s_2 / (s_1 + s_2). \quad (6.11)$$

Тогда уравнение (6.9) можно представить в виде

$$\Delta p = -(s_1 + s_2)(p - \hat{p})p(1-p)/\bar{w}. \quad (6.12)$$

Отсюда легко видеть, что если $p > \hat{p}$, то Δp отрицательно, но если $p < \hat{p}$, то Δp положительно. Это означает, что если частота гена отклонилась от точки равновесия, то она будет стремиться вернуться к этой точке. Иначе говоря, имеет место устойчивое равновесие. С тех пор как это положение было продемонстрировано Фишером [105], многие генетики рассматривали сверхдоминирование как главный фактор, поддерживающий генетическую изменчивость внутри вида. Однако примеры сверхдоминирования, или однолокусного гетерозиса по приспособленности, весьма немногочисленны. Одним из таких широко известных примеров является серповидноклеточная анемия у человека.

Мутантный гемоглобин Hb^S , ответственный за серповидноклеточную анемию, отличается от нормального гемоглобина Hb^A тем, что в положение 6 β -цепи вместо глутаминовой кислоты находится валин [192]. Аллель Hb^S селективно невыгоден, поскольку обуславливает тяжелую анемию у гомозигот, однако у гетерозигот Hb^SHb^A он получает преимущество, по-

сколько обеспечивает устойчивость к малярии [4]. Используя терминологию, введенную для фенотипического отбора, мы можем назвать отбор, поддерживающий равновесие благодаря сверхдоминированию, стабилизирующим (на генотипическом уровне). Заметим, что ω , являясь функцией p , достигает максимума при $\hat{p}$, задаваемом выражением (6.11). Поэтому равновесие при сверхдоминировании находится в точке локального максимума функции приспособленности.

В противоположной ситуации, когда обе гомозиготы имеют преимущество перед гетерозиготой ($\omega_1 > 1$ и $\omega_2 > 1$), равновесная частота также задается выражением (6.11), но равновесие неустойчиво. В этом можно сразу убедиться, обратившись к формуле (6.12): если $p > \hat{p}$, то Δp положительно, если же $p < \hat{p}$, то Δp отрицательно. В любом случае отклонения от равновесия будут возрастать в каждом последующем поколении. Следовательно, если частота какого-то аллеля ниже его равновесной величины, то аллель в конце концов будет утрачен популяцией, а в противном случае он в конечном счете зафиксируется. К такому случаю невыгодности гетерозигот (отрицательному сверхдоминированию) можно применить термин *дизруптивный* (или *центробежный*) отбор.

Тип частотно-зависимого отбора, при котором каждый аллель, становясь редким, получает преимущество («преимущество меньшинства»), долгое время рассматривали как важный механизм поддержания внутривидовой генетической изменчивости [54, 161, 597]. Такой отбор будет действовать, в частности, когда среди разных генотипов существует «разделение труда».

В качестве простой модели частотно-зависимого отбора предположим, вслед за Райтом и Добржанским [605], что относительные приспособленности A_1A_1 , A_1A_2 и A_2A_2 равны $1+a-bp$, 1 и $1-a+bp$, a и b — положительные константы ($b > a > 0$), а p — частота аллеля A_1 (см. табл. 6.1). При случайном скрещивании изменение частоты гена за поколение под действием отбора равно

$$\Delta p = p(1-p)(a-bp)/\bar{\omega}, \quad (6.13)$$

где

$$\bar{\omega} = 1 - (a-bp)(1-2p). \quad (6.13a)$$

В этой модели частота A_1 будет возрастать, если $p < a/b$, и уменьшаться, если $p > a/b$. Ясно, что $\hat{p} = a/b$ является точкой устойчивого равновесия. Более того, в этой точке все три генотипа имеют одинаковую приспособленность. Однако в общем случае частотно-зависимого отбора это условие не обязательно должно выполняться. Оно имеет место только в том случае,

Таблица 6.1. Простая модель частотно-зависимого отбора

Генотип	A_1A_1	A_1A_2	A_2A_2
Приспособленность	$1+a-bp$	1	$1-a+bp$
Частота	p^2	$2p(1-p)$	$(1-p)^2$

когда гетерозигота по приспособленности занимает в точности промежуточное положение между двумя гомозиготами, и фактически утрачивается при несколько более общей форме частотно-зависимого отбора, который я рассмотрю позже (табл. 6.2).

6.4. Другие виды терминологии и концепции, использующиеся при анализе естественного отбора

Применительно к такому типу отбора, который порождает генетическую неоднородность или полиморфизм, Добржанским [85] был предложен термин «балансирующий отбор». Он включает сверхдоминирование и частотно-зависимый отбор с преимуществом меньшинства, о чем мы уже говорили выше. В качестве одной из форм уравнивающего отбора Добржанский рассматривает «разнообразящий отбор», который, по его мнению, ответствен за некоторые формы сбалансированного полиморфизма. Согласно Добржанскому, этот термин является синонимом дизруптивного отбора (см. [86], с. 127) и применим к тем случаям, когда в разных экологических нишах предпочтительными оказываются разные генотипы. Однако я должен заметить, что этот тип отбора необязательно ведет к сбалансированному полиморфизму и часто для его поддержания необходимы специальные условия (см., например, [350]).

Временные изменения интенсивности и направления естественного отбора в зависимости от ситуации либо ускоряют, либо замедляют фиксацию гена (поддерживая полиморфизм). Элегантная модель такого процесса рассмотрена в работе [167]. Пусть A и a — пара аллелей по аутосомному локусу, причем аллель a полностью рецессивен по приспособленности. Предполагая случайное скрещивание, авторы показали, что в реальных с биологической точки зрения условиях полиморфизм будет поддерживаться в том случае, когда среднеарифметическое приспособленности рецессивного аллеля по отношению к доминантному аллелю больше единицы, а соответствующее среднегеометрическое — меньше единицы. Например, это условие может выполняться, если приспособленность рецессивного аллеля в норме на 5—10% выше, чем доминантного, но в результате

эпидемического заболевания через каждые двадцать поколений все рецессивы погибают. Во время каждой эпидемии частота рецессивного гена резко уменьшается, затем постепенно увеличивается вплоть до следующей эпидемической вспышки. Более подробно эта проблема рассмотрена в работе [211]. Отметим, что нецелесообразно причислять к категории дизруптивного отбора случаи, при которых коэффициенты отбора флуктуируют во времени (например, когда в разных поколениях оказываются предпочтительными разные аллели или генотипы).

Остановимся также на терминологии, предложенной для того, чтобы осветить разные аспекты действия естественного отбора. При описании подхода, который применяют для определения доли особей, оставляемой для размножения, Уоллес ввел термины «жесткий» и «мягкий» отбор [564]. Их смысл проще всего уяснить, рассмотрев отсекающий отбор, представленный графиком *B* на рис. 6.2, где особей со значением фенотипического признака, превышающим значение в «граничной точке» *C*, оставляют для размножения, а остальных элиминируют. При жестком отборе эта точка фиксирована, и, таким образом, доля популяции, которая выживает и размножается, колеблется, так как положение частотного распределения (штриховая кривая) непостоянно: если распределение сдвигается влево, доля выживших особей уменьшается, если вправо, то увеличивается. При мягком отборе граничная точка может сдвигаться с изменением положения кривой распределения таким образом, что доля выживших особей остается постоянной. Следовательно, при мягком отборе средняя приспособленность популяции сохраняется постоянной, даже если генетический состав ее сильно изменяется. Такой тип отбора неприменим в случае рецессивных леталей, которые в гомозиготном состоянии обязательно ведут к гибели. Однако если признать, что слабые конкуренты все же могут выжить в отсутствие сильных, то станет ясно, что мягкий отбор встречается в природе не так уж редко.

Мак-Артур и Уилсон [318] в своей хорошо известной работе по островной биогеографии рассматривают два типа естественного отбора, которые они называют *r*-отбором и *K*-отбором. Обозначим численность популяции через *N* и предположим, что скорость ее изменения описывается уравнением

$$dN/dt = rN(K - N)/K, \quad (6.14)$$

где *r* — скорость свободного роста популяции, а *K* — «предельная плотность насыщения» (максимальная численность). Если $N \ll K$, то размер популяции увеличивается экспоненциально со скоростью *r*, но если $N \rightarrow K$, то скорость роста уменьшается до тех пор, пока не будет достигнута численность $N = K$. Уравнение (6.14) часто называют логистическим. В условиях невысо-

кой численности, когда среда «не загружена», приспособленность определяется в основном величиной r , поэтому преимущество приобретают такие типы особей, которые потребляют наибольшее количество пищи, хотя это и расточительно. Такую ситуацию называли r -отбором; эволюция предпочитает здесь «продуктивность». С другой стороны, если численность в данных условиях слишком велика, то большее значение приобретает предельная плотность насыщения популяции. Теперь выживают генотипы с меньшими потребностями в пище. Такая ситуация названа K -отбором. Здесь эволюция предпочитает, так сказать, «эффективность преобразования пищи в потомство». Концепция K - и r -типов отбора получила широкое признание среди экологов. Например, Рафгарден [495], предположив наличие некой связи между параметрами r и K , обсуждает поведение популяции, содержащей пару аллелей с разными значениями r и K . В целом, однако, мне кажется, что эту концепцию подчас применяли без достаточных на то оснований.

Итак, мы рассмотрели различные типы отбора и обсудили их роль; перейдем теперь к более трудному вопросу — оценке интенсивности отбора. Особенно важно количественно оценить действие отбора на индивидуальные локусы, когда известна полная интенсивность фенотипического отбора.

6.5. Генетический груз

Рассматривая генетический груз в качестве меры интенсивности генотипического отбора, Кроу [59] определяет эту величину как долю, которую составляет разность между средней приспособленностью популяции и приспособленностью оптимального (наиболее приспособленного) генотипа от приспособленности последнего. В аналитическом виде это определение можно записать как

$$L = \frac{\omega_{\text{опт}} - \bar{\omega}}{\omega_{\text{опт}}}, \quad (6.15)$$

где $\bar{\omega}$ — средняя приспособленность популяции, а $\omega_{\text{опт}}$ — приспособленность оптимального генотипа. Рассмотрим вкратце некоторые наиболее интересные результаты, полученные при исследовании генетического груза (подробности читатель может найти в работах [67], с. 297 и [257], с. 447).

Мутационный груз

Появление этого груза обусловлено элиминацией повторных вредных мутаций. Рассмотрим некий аутосомный локус в очень большой диплоидной популяции со случайным скрещиванием. Холдейн [158] впервые показал, что снижение приспособлен-

ности, обусловленное появлением полностью рецессивных вредных мутаций, равно скорости мутирования (на гамету), т. е. в этом случае $L_{\text{мут}} = v$. Для полудоминантных мутаций $L_{\text{мут}} = 2v$, т. е. груз равен скорости мутирования на зиготу. Существенно, что в этих случаях груз почти не зависит от степени вредности мутаций. Точно такой же вывод независимым путем был получен и Мёллером [388], который использовал концепцию «генетической гибели». Вообще, если приспособленность мутантных гетерозигот или гомозигот равна $1 - hs$ и $1 - s$ по отношению к приспособленности нормальной гомозиготы (которая принимается за единицу), $L_{\text{мут}}$ изменяется от v до $2v$ при изменении h от 0 до 1. Правда, можно показать, что $L_{\text{мут}}$ очень близок к $2v$, если $sh^2 \gg 2v(|1 - 2h|)$. Поскольку для летальных и полублетальных генов скорость мутирования и степень доминирования равны примерно $v = 10^{-5}$ и $h \approx 0,02$, то формула

$$L_{\text{мут}} = 2v \quad (6.16)$$

с достаточной степенью точности позволяет оценить мутационный груз в природных популяциях. Другими словами, мутационный груз равен скорости мутирования на зиготу. Кроу [58] назвал это правило принципом Холдейна — Мёллера.

Мутационный груз можно представить и как функцию s , h и v [226]. Можно также показать, что мутационный груз не меняется, даже если принять во внимание регуляцию численности популяции ([257], с. 48). Однако существует целый ряд условий, при которых принцип Холдейна — Мёллера должен быть несколько модифицирован. К наиболее важным относится случай эпистаза «усиленного типа» между локусами, когда наличие двух и более мутантных генов у одной особи оказывает более вредный эффект, чем это можно ожидать при простом суммировании вредных эффектов мутаций. Кимура и Маруяма [249] показали, что если вредный эффект пропорционален квадрату числа мутантных генов, то генетический груз в условиях случайного скрещивания достигает примерно половины той его величины, которая бывает в отсутствие эпистаза. Следует ожидать, что по мере возрастания «усиленного» эпистаза груз должен уменьшаться. Позднее Кроу и Кимура [69] развили теорию, которая позволяет рассчитать эффективность отсекающего отбора, и применили ее к мутантам по жизнеспособности у *Drosophila*, показав, что генетический груз при этом действительно становится меньше. Еще раньше Мукаи [372] установил, что мутации, вызывающие небольшое снижение жизнеспособности гомозигот (названные им полигенами жизнеспособности), встречаются гораздо чаще, чем летали и полублетали, составляя около 0,4 гаметы на поколение. Каждая из таких мутаций снижает жизнеспособность в гомозиготном со-

стоянии в среднем примерно на 2—3%. Они имеют гораздо большую степень доминирования, чем летали (с $h \approx 0,4$ для вновь возникших мутаций), причем эффекты нескольких локусов примерно аддитивны. Если исходить из принципа Холдейна — Мёллера, то мутационный груз, обусловленный только полигенами жизнеспособности, должен достигать 0,8. Однако при отсекающем отборе (или допустимых отклонениях от него, проиллюстрированных на рис. 6.2, *А* и *Б*) груз может быть значительно меньше. Примем за основное фенотипическое значение X плейотропные гетерозиготные эффекты мутаций в отношении жизнеспособности. Такой выбор основан на том, что эти мутации существуют главным образом в гетерозиготном состоянии (в связи с их низкой частотой), а их влияние на приспособленность обуславливается изменением не только жизнеспособности, но и плодовитости. Итак, X — это шкала действия отсекающего отбора; она соответствует «потенциалу приспособленности» Милкмена [359]. Коэффициент отбора s против такой мутации и стандартное отклонение σ потенциала приспособленности можно оценить экспериментально: $s \approx -0,02$, а $\sigma = 0,02 \sim 0,04$. Обозначим через A средний эффект единичной мутации в отношении потенциала приспособленности и допустим, что имеет место 10%-ный ограничивающий отбор ($\bar{w} = 0,9$). Тогда можно показать, что $A = -0,002 \sim -0,005$. Это означает, что при ограничивающем отборе мутации с $(0,1 \sim \approx 0,4)$ %-ным эффектом в отношении приспособленности элиминируются из популяции столь же эффективно, как если бы они обладали 2%-ным эффектом, но подвергались действию аддитивного отбора. Отсекающий отбор элиминирует мутации группами, так что груз становится гораздо меньше, чем ожидается исходя из принципа Холдейна — Мёллера. Отсекающий отбор в строгом смысле этого слова вряд ли встречается в природе; однако вполне может существовать так называемый *квазиотсекающий отбор* (рис. 6.2, *А* и *Б*), который обладает эффективностью, не слишком отличающейся от эффективности отсекающего отбора.

В случае малых популяций, где для частот мутаций характерна тенденция к дрейфу от равновесных значений, принцип Холдейна — Мёллера также необходимо модифицировать. Эту проблему исследовали Кимура, Маруяма и Кроу [252], которые показали, что в очень малых популяциях груз значительно выше, чем в больших.

Сегрегационный груз

Такой груз возникает в том случае, когда оптимальный генотип не может быть единственным в популяции. Самым интересным примером такого рода является сверхдоминирование. Рассмотрим популяцию со случайным скрещиванием и предпо-

ложим, что аллели A_1 и A_2 имеют частоты p и $q (=1-p)$. Пусть $1-s_1$, 1 и $1-s_2$ — относительные приспособленности генотипов A_1A_1 , A_1A_2 и A_2A_2 соответственно, так что s_1 и s_2 — коэффициенты отбора против гомозигот A_1A_1 и A_2A_2 . Выше было показано [см. формулы (6.9—6.12)], что равновесная частота $p = s_2/(s_1+s_2)$ устойчива, если $s_1 > 0$ и $s_2 > 0$. Тогда, подставив ее в (6.10) и заметив, что $w_{\text{опт}} = 1$, получим выражение для сегрегационного груза

$$L_c = s_1 s_2 / (s_1 + s_2). \quad (6.17)$$

Так как обычно считается, что коэффициенты отбора s_1 и s_2 гораздо больше скоростей мутирования, то сегрегационный груз, обусловленный сверхдоминированием, должен быть гораздо большим в расчете на локус, чем соответствующий мутационный груз. Если имеются множественные аллели $A_1, A_2, \dots, A_n$, причем все гетерозиготы имеют одинаковую приспособленность, а относительная приспособленность гомозиготы $A_i A_i$ (по сравнению с приспособленностью гетерозигот) равна $1-s_i$ ($i=1, 2, \dots, n$), то сегрегационный груз становится равным $L_c = \bar{s}/n$, где $\bar{s}$ — среднегармоническое для n коэффициентов отбора (см. [67], с. 305; [257], с. 56). Тогда, если n велико, груз будет малым. Однако в конечной популяции набор аллелей ограничен случайным дрейфом, даже если в результате мутаций продуцируется бесконечное число аллелей со сверхдоминированием. Кимура и Кроу [246] оценили эффективное число сверхдоминантных аллелей при данных скорости мутирования и эффективном размере популяции, используя метод диффузионных уравнений и предположив, что процесс мутирования всегда приводит к возникновению какого-то нового, не существовавшего раньше, сверхдоминантного аллеля. Они пришли к выводу, что сверхдоминирование является не очень эффективным фактором для поддержания большого числа аллелей на локус в конечной популяции и что обусловленный сверхдоминированием груз (на локус) становится при этом весьма значительным. Следовательно, ситуация, при которой очень большое число сегрегирующих локусов поддерживается за счет значительного сверхдоминантного эффекта каждого из них (в предположении независимого мультипликативного эффекта индивидуальных локусов в отношении общей приспособленности), маловероятна. Чтобы обойти трудности, связанные с большим сегрегационным грузом, Кинг [267], Свед, Рид и Бодмер [539], а также Милкмэн [356] предложили модель с отсекающим отбором (см. гл. 2). Фактически такие же типы моделей, но только более детальные Свед [538], Милкмэн [359] и Уиллз [581] привлекли для объяснения белкового полиморфизма. Эти авторы предполагают, что степень приспособленности пропорциональна числу гетерозигот-

ных локусов, и утверждают, что отсекающий отбор или какая-то разумная его модификация совместно с «мягким» отбором не приводят к большому грузу и поддерживают полиморфизм по тысячам локусов. Уиллз [581], в частности, для обозначения такого типа естественного отбора использует термин «ранжированный отбор». Он предполагает, что потенциал приспособленности прямо пропорционален средней гетерозиготности, а сама приспособленность определяется соотношением между «рангом» приспособленности данного организма и суровостью внешней среды. Он утверждает, что «ранжированный» отбор способен поддерживать самые разнообразные типы генетического полиморфизма. Согласно этому утверждению, число типов полиморфизма, которые могут сохраняться в условиях определенного селективного равновесия, возрастает пропорционально квадрату численности популяции. По-моему, нет никаких оснований полагать, как это делает Уиллз, что ранжированный отбор, основанный исключительно (или главным образом) на числе гетерозиготных сайтов в геноме, включая даже полиморфизм по длине нетранскрибируемых участков генов рРНК у *Xenopus* [480], действует постоянно, формируя генетическую структуру природных популяций. Такой взгляд должен опираться на безусловную веру в преимущество гетерозиготности *per se*. Веским доводом против него наряду с другими фактами является очень малая величина дисперсии сверхдоминирования по сравнению с аддитивной генетической дисперсией жизнеспособности в разных популяциях *Drosophila melanogaster* [373] (см. также разд. 9.2). Кроме того, теоретические результаты, говорящие о том, что уровни разных типов полиморфизма возрастают как квадрат численности популяции, просто не согласуются с наблюдениями (см., например, [523]).

Генетический груз при частотно-зависимом отборе

Нередко утверждают (см., например, [548]), что частотно-зависимый отбор в пользу редких генотипов не сопряжен ни с каким генетическим грузом. Чтобы оценить, насколько верно такое утверждение, рассмотрим простую модель частотно-зависимого отбора (см. табл. 6.2), где s_1 , s_2 , c_1 и c_2 — положительные константы, а p и q — частоты аллелей A_1 и A_2 . Эта модель [225] рассматривает такую ситуацию, когда условия местообитания данного вида подразделяются на два типа: E_1 и E_2 ; особи A_1A_1 хорошо приспособлены к E_1 , но не к E_2 , а особи A_2A_2 , наоборот, хорошо приспособлены к E_2 , но не к E_1 . Предполагается также, что гетерозиготные особи A_1A_2 одинаково приспособлены к обоим средам. Тогда при случайном скрещивании средняя приспособленность популяции равна

$$\bar{w} = 1 - s_1 p^2 (p - c_1) - s_2 q^2 (q - c_2), \quad (6.18)$$

Таблица 6.2. Модель частотно-зависимого отбора

Генотип	A_1A_1	A_1A_2	A_2A_2
Приспособленность	$1 - s^1(p - c_1)$	1	$1 - s_2(q - c_2)$
Частота	p^2	$2pq$	q^2

где $q = 1 - p$. Равновесие достигается при равных средних приспособленностях A_1 и A_2 и поэтому определяется из формулы

$$s_1 p (p - c_1) = s_2 q (q - c_2). \quad (6.19)$$

Для простоты в дальнейшем будем считать, что коэффициенты s_1 и s_2 равны ($s_1 = s_2$), и обозначим их через s . Из уравнения (6.19) легко найти равновесную частоту

$$\hat{p} = (1 - c_2) / (2 - c_1 - c_2). \quad (6.20)$$

Далее всюду крышка ($\sim$) над символами будет означать, что соответствующие величины являются равновесными. Изменение частоты A_1 за одно поколение равно

$$\Delta p = spq [(1 - c_2) - (2 - c_1 - c_2) p] / \bar{w}, \quad (6.21)$$

где

$$\bar{w} = 1 - sp^2 (p - c_1) - sq^2 (q - c_2). \quad (6.22)$$

Как видно из соотношения (6.21), для того чтобы это равновесие было устойчивым, должно выполняться неравенство $2 - c_1 - c_2 > 0$. Кроме того, для существования нетривиального равновесия необходимо, чтобы $0 < \hat{p} < 1$ и $0 < 1 - \hat{p} < 1$. Объединяя эти неравенства, получаем необходимое и достаточное условие существования нетривиального устойчивого равновесия:

$$\left. \begin{aligned} 0 < c_1 < 1 \\ 0 < c_2 < 1 \end{aligned} \right\}. \quad (6.23)$$

Тогда относительные приспособленности A_1A_1 и A_2A_2 при равновесии есть

$$\hat{w}_{11} = 1 - s(1 - c_1)(1 - c_1 - c_2) / (2 - c_1 - c_2), \quad (6.24)$$

$$\hat{w}_{22} = 1 - s(1 - c_2)(1 - c_1 - c_2) / (2 - c_1 - c_2). \quad (6.25)$$

Средняя приспособленность популяции при этом равна

$$\hat{\bar{w}} = 1 - s(1 - c_1)(1 - c_2)(1 - c_1 - c_2) / (2 - c_1 - c_2)^2. \quad (6.26)$$

Рассмотрим теперь следующие три случая. Во-первых, если $c_1 + c_2 = 1$, то все генотипы имеют при равновесии одинаковую

приспособленность ($\hat{w}_{11}=1=\hat{w}_{22}$) и, значит, груз отсутствует. Это соответствует рассмотренному в предыдущем разделе случаю [см. формулы (6.13) и (6.13a)]. Во-вторых, если $0 < c_1 + c_2 < 1$, то $\hat{w}_{11} < 1$ и $\hat{w}_{22} < 1$, так что гетерозигота имеет наибольшую приспособленность. Это означает, что A_1 и A_2 при равновесии характеризуются сверхдоминированием. Тогда возникает генетический груз, равный

$$L = s(1 - c_1)(1 - c_2)(1 - c_1 - c_2)/(2 - c_1 - c_2)^2. \quad (6.27)$$

В-третьих, если $1 < c_1 + c_2 < 2$, то $\hat{w}_{11} > 1$ и $\hat{w}_{22} > 2$, так что обе гомозиготы более приспособлены, чем гетерозигота. Замечательно, что в этом случае, в отличие от случая отрицательного сверхдоминирования с постоянной приспособленностью, равновесие является устойчивым. Предположим теперь, что $c_1 > c_2$, так что наиболее приспособленным среди трех генотипов является генотип A_2A_2 . Тогда груз равен

$$L = s[(1 - c_2)^2(c_1 + c_2 - 1)/(2 - c_1 - c_2)^2]/\hat{w}_{22}. \quad (6.28)$$

Интересно также, что по сравнению со случаем постоянной приспособленности частота гена в точке устойчивого равновесия и частота, при которой достигается максимум $\bar{w}$, как правило, неодинаковы. Фактически частота A_1 , которая максимизирует $\bar{w}$ [см. соотношение (6.22)], равна

$$\hat{p} = (3 - 2c_2)/(6 - 2c_1 - 2c_2), \quad (6.29)$$

что отличается от $\hat{p}$, задаваемого (6.20).

Дрейф и дисметрический груз

Как мы уже видели, в рамках приведенной выше модели частотно-зависимого отбора (табл. 6.2) в случае, когда $s_1 = s_2$ и $c_1 + c_2 = 1$, все три генотипа при устойчивом равновесии имеют одинаковую приспособленность и, следовательно, генетический груз отсутствует. В этом случае модель сводится к ситуации, описываемой в табл. 6.1, которая означает, что приспособленность гетерозиготы предполагается всегда строго промежуточной между приспособленностями двух гомозигот (случай «без доминирования»). Однако трудно представить, чтобы даже в такой упрощенной ситуации не было груза; очевидно, в природе под влиянием дрейфа частота гена отклоняется от своего значения в точке устойчивого равновесия и появляется груз.

В общем случае мы можем определить обусловленный дрейфом груз как

$$L_{\text{дрейф}} = \left[\hat{\bar{w}} - E(\bar{w}) \right] / \hat{\bar{w}}, \quad (6.30)$$

где $\bar{\omega}$ — средняя приспособленность в бесконечно большой популяции при равновесии, а $E(\bar{\omega})$ — ожидаемая величина средней приспособленности в конечной популяции. Кимура и Ота [255] показали, что в приведенной выше модели частотно-зависимого отбора (табл. 6.1) генетический груз, обусловленный дрейфом, в конечной популяции с эффективной численностью N_e равен

$$L_{\text{дрейф}} = 1/(2N_e). \quad (6.31)$$

Примечательно, что этот груз не зависит от параметров a и b , а определяется только эффективным размером популяции. Сходное свойство такого груза несколько ранее было описано Робертсоном [490], который показал, что если в популяции с эффективной численностью N_e сегрегирует пара сверхдоминантных аллелей, груз $L_{\text{дрейф}} = 1/(4N_e)$ не зависит от коэффициентов отбора.

Одно из важных свойств частотно-зависимого отбора, о котором мы уже говорили, состоит в том, что частота гена в состоянии устойчивого равновесия ($\hat{p}$) обычно не равна той частоте ($\bar{p}$), при которой достигается максимум средней приспособленности $\bar{\omega}$ [ср. соотношения (6.20) и (6.29)]. Эту проблему мы рассмотрим несколько позже в рамках модели, представленной в табл. 6.1. В данной модели равновесная частота гена $\bar{p} = a/b$, а соответствующая средняя приспособленность $\bar{\omega} = 1$. С другой стороны, частота гена, которая максимизирует $\bar{\omega}$, есть $\bar{p} = 1/4 + a/(2b)$, и соответствующая максимальная приспособленность равна $\bar{\omega}_{\text{max}} = 1 + (2a - b)^2/(8b)$. Это означает, что естественный отбор поддерживает приспособленность популяции на уровне, который немного ниже возможной максимальной величины. Таким образом, этот дефицит приспособленности популяции можно рассматривать как груз. Первым, кто отметил этот факт, был, по-видимому, Холдейн. В личном сообщении Кроу (1959, не опубликовано) Холдейн предложил называть этот тип груза «дисметрическим грузом». Поэтому формально мы определим дисметрический груз как

$$L_{\text{дис}} = (\bar{\omega}_{\text{max}} - \bar{\omega}) / \bar{\omega}_{\text{max}} \quad (6.32)$$

В рамках упоминавшейся модели частотно-зависимого отбора мы имели $L_{\text{дис}} = (2a - b)^2/[8b + (2a - b)^2]$. Используя такую же модель для анализа полиморфизма, связанного с ST- и CN-хромосомами у *Drosophila pseudoobscura*, Райт и Добжанский [605] получили следующие оценки: $a = 0,902$; $b = 1,288$. Таким образом, для данного случая $L_{\text{дис}} = 0,0252$, или около 2,5%.

Субституционный груз

В процессе замены одного аллеля другим под действием естественного отбора (в связи с изменением среды) происходит элиминация менее благоприятного аллеля и, следовательно, возникает генетический груз. Концепция субституционного груза принадлежит Холдейну [163], который назвал его «ценой естественного отбора». Замечательное свойство этого груза состоит в том, что он почти не зависит от коэффициента отбора, а определяется главным образом начальной частотой мутантного аллеля. Данный вопрос я уже обсуждал подробно в гл. 2 (разд. 2.2) и здесь лишь отмечу, что эта концепция все еще остается спорной. Большинство авторов, высказывающихся против, говорят о том, что замену менее благоприятного аллеля более благоприятным нельзя расценивать как груз, так как приспособленность вида при этом повышается. По моему, здесь не учитывается тот важный факт, что как физическая, так и биотическая среда для каждого вида периодически становятся неблагоприятными, тогда как появление мутаций, благоприятных уже «на старте», — всегда очень редкое событие. Однако не исключено, что при некоторых условиях величина груза, или «цены», может быть меньше, чем предусматривается принципом «цены естественного отбора» Холдейна. Ней обсуждает эту проблему, оценивая, какой должна быть плодовитость, необходимая для замещения гена ([401], с. 61—66). Ряд вопросов, касающихся «цены» или субституционного груза, рассматриваются у Кроу [63], Кроу и Кимуры ([67], с. 244—252), Фельзенштейна [101] и Кимуры и Оты ([257], с. 72—88).

6.6. Модель процесса адаптации, предложенная Фишером

Для того, чтобы организм или какой-то орган могли хорошо адаптироваться к данной среде, необходимо, чтобы к ней были хорошо адаптированы сразу многие сложные признаки. Фишер [107] для пояснения этого положения привлекает в качестве аналогии юстировку микроскопа. Эта процедура состоит из небольших, независимых друг от друга изменений положения разных частей прибора, например смещения линз по вертикали или горизонтали, поворотов их под разными углами, изменяющих показатель преломления, и т. д. Фишер высказал мысль, что в более абстрактной форме процесс адаптации можно представить геометрически, так, как это сделано для двумерного случая на рис. 6.4. Здесь степень адаптации, или соответствия, измеряется расстоянием точки *A* (отвечающей состоянию популяции в данный момент) до фиксированной точки *O*, соответствующей наиболее приспособленному, или оп-

тимальному, состоянию. Любая точка внутри окружности, проходящей через точку A и имеющей центр в точке O , отвечает состоянию более высокой адаптации. При смещении A на расстояние r в каком-либо направлении (в результате мутации или рекомбинации) степень адаптации может увеличиваться, если точка смещается ближе к центру окружности, или уменьшаться, если смещение направлено от центра. Легко видеть, что если r очень мало, то вероятности этих двух событий примерно равны. Это означает, что вероятность повышения адаптации

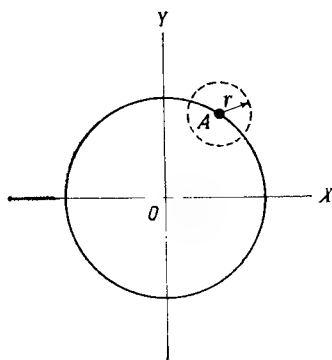


Рис. 6.4. Иллюстрация модели адаптивного процесса Фишера в двумерном пространстве.

приближается к предельному значению $1/2$, когда r стремится к нулю. С другой стороны, если r больше диаметра окружности d , то вероятность повышения адаптации равна нулю. Такие же рассуждения можно провести и для трехмерного случая, когда, согласно Фишеру, эта вероятность равна $(1-r/d)/2$ при $0 < r \leq d$. Попробуем теперь найти вероятность повышения адаптации для многомерного случая. Согласно Фишеру, при большом n искомая вероятность равна

$$p = \frac{1}{\sqrt{2\pi}} \int_x^{\infty} e^{-\lambda^2/2} d\lambda, \quad (6.33)$$

где $x = r\sqrt{n/d}$.

Я решил эту задачу, обратившись к многомерной геометрии, и получил, что для $0 \leq r/d \leq 1$ вероятность повышения адаптации равна

$$p = \frac{1}{2} I_z(v/2, 1/2), \quad (6.34)$$

где $z = 1 - (r/d)^2$, $v = n - 1$. В этой формуле функция $I_z(\cdot, \cdot)$ задается отношением

$$I_z(a, b) = B_z(a, b) / B(a, b), \quad (6.35)$$

где

$$B_z(a, b) = \int_0^z x^{a-1} (1-x)^{b-1} dx, \quad (6.36)$$

— это неполная бета-функция первого рода, а $B(a, b)$ — бета-функция Эйлера. Далее можно показать, что если $v=n-1$ велико, то формулу (6.34) можно аппроксимировать выражением

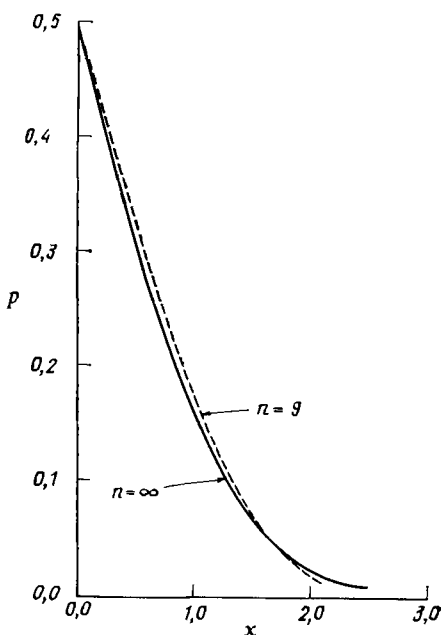


Рис. 6.5. Зависимость между вероятностью повышения адаптивности p (ордината) и величиной изменения x (абсцисса). Здесь $x=r\sqrt{n/d}$, а n — размерность пространства.

(6.33). Соответствующая формула для $I_z(\cdot, \cdot)$ есть в справочнике Абрамовица и Стигана [1]; неполная бета-функция затабулирована Пирсоном [465]. На рис. 6.5 приведены зависимости вероятности повышения адаптивности p от величины изменения x для $n=\infty$ (сплошная кривая) и $n=9$ (штриховая кривая). Заметим, что формула Фишера (6.33) дает достаточно хорошее приближение уже для 9-мерного пространства.

6.7. Зависимость между фенотипическим и генотипическим отбором

Общая формулировка

Чтобы оценить, насколько велика роль случайного дрейфа генов в процессе эволюции, необходимо знать, какова интенсивность того отбора, который индуцируется на уровне аллелей

отдельных локусов при данной интенсивности фенотипического отбора. Рассмотрим какой-либо количественный признак, например рост, вес, концентрацию какого-то вещества или более абстрактный количественный признак, достаточно репрезентативный для оценки дарвиновской приспособленности. Допустим, что этот признак детерминируется большим числом локусов, каждый из которых обладает очень малым фенотипическим эффектом и к тому же подвержен влиянию среды (полу-

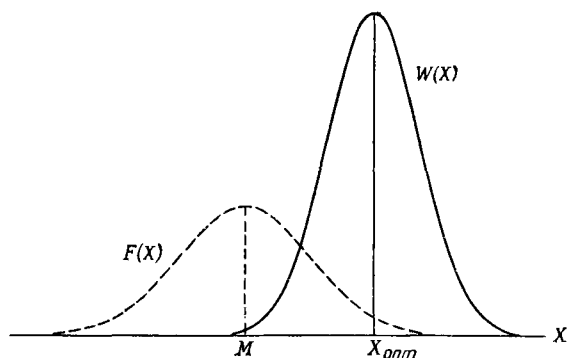


Рис. 6.6. Примеры функций частоты и приспособленности. Функция частоты (штриховая кривая) имеет среднее значение M и дисперсию σ^2 . Функция приспособленности (сплошная кривая) достигает максимума в точке $X=X_{\text{опт}}$.

чаемые результаты будут такими же, если мы рассмотрим более мелкие генетические единицы, например нуклеотиды). Чтобы упростить анализ, предположим также, что соответствующие гены аддитивны в отношении данного признака. Используем метод Кимуры и Кроу [247] и обозначим через X измеряемый фенотипический признак со средним значением M и дисперсией σ^2 , а через $X_{\text{опт}}$ — оптимальное значение этого признака. Пусть $F(X)$ и $W(X)$ — относительная частота и приспособленность особей с величиной признака X (см. рис. 6.6).

Рассмотрим отдельный локус, в котором пара аллелей A_1 и A_2 сегрегирует с частотами $1-p$ и p соответственно. Предположим, что имеется диплоидная популяция со случайным скрещиванием, в которой X_{ij} — среднее значение признака у особей $A_i A_j$, где $i=1$ или 2 и $j=1$ или 2 ($X_{12}=X_{21}$).

Часто бывает удобно измерять параметры, характеризующие величину какого-либо признака, в единицах фенотипического стандартного отклонения σ . Для этих целей при обозначениях мы будем использовать строчные буквы. Кроме того, если это специально не оговорено, мы будем брать $X_{\text{опт}}$ в качестве точки отсчета и рассматривать не исходные значения признака, а их нормированные отклонения $x=(X-X_{\text{опт}})/\sigma$. Соот-

ветствующие функции частоты и приспособленности обозначим через $f(x)$ и $w(x)$. Пусть также $m = (M - X_{\text{опт}})/\sigma$ и $a_{ij} = (X_{ij} - M)/\sigma$. Заметим, что a_{ij} есть отклонение значения рассматриваемого признака для $A_i A_j$ от популяционного среднего в единицах σ .

Предположим, что характер распределения признака одинаков для разных генотипов по этому локусу и с хорошим приближением описывается функцией $f(x)$, поскольку предполагается, что индивидуальные генные эффекты пренебрежимо ма-

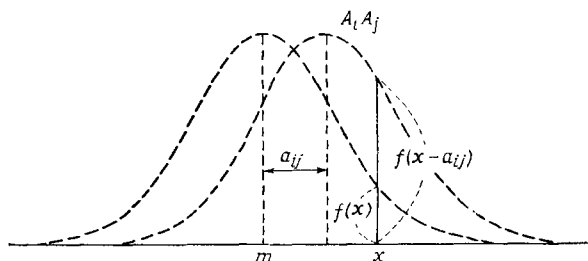


Рис. 6.7. Схема, иллюстрирующая смысл уравнения (6.37). Здесь a_{ij} — отклонение среднего значения фенотипического признака особей $A_i A_j$ от популяционного среднего в единицах σ .

лы. Пусть w_{ij} — относительная приспособленность особей $A_i A_j$; тогда

$$w_{ij} = \int_{-\infty}^{\infty} w(x) f(x - a_{ij}) dx, \quad (6.37)$$

(см. рис. 6.7). Предположив, что a_{ij} бесконечно мало, разложим $f(x - a_{ij})$ в ряд Тейлора:

$$f(x - a_{ij}) = f(x) - a_{ij} f'(x) + (a_{ij}^2/2) f''(x) - \dots \quad (6.38)$$

Полностью это разложение проведено в работе Кимуры и Кроу [247], но мы ограничимся здесь членом второго порядка малости. Тогда из уравнения (6.37) получим

$$w_{ij} = b_0 - a_{ij} b_1 + a_{ij}^2 b_2/2, \quad (6.39)$$

где

$$b_0 = \int_{-\infty}^{\infty} w(x) f(x) dx, \quad (6.39a)$$

$$b_1 = \int_{-\infty}^{\infty} w(x) f'(x) dx, \quad (6.39b)$$

$$b_2 = \int_{-\infty}^{\infty} w(x) f''(x) dx. \quad (6.39b)$$

Здесь знак «штрих» означает дифференцирование.

Обозначим через a тот эффект, который оказывает на величину признака x замена A_1 на A_2 . При случайном скрещивании частоты трех генотипов A_1A_1 , A_1A_2 и A_2A_2 равны соответственно $(1-p)^2$, $2(1-p)p$ и p^2 , где p — частота A_2 . Предположив, что эффекты генов в отношении величины x аддитивны, нетрудно показать, что $a_{11} = -2ap$, $a_{12} = a(1-2p)$, $a_{22} = 2a(1-p)$. Тогда, используя уравнение (6.39), можно найти среднюю приспособленность популяции: $\bar{w} = w_{11}(1-p)^2 + w_{12}2(1-p)p + w_{22}p^2$, и среднюю приспособленность аллеля A_2 : $w_2 = w_{21}(1-p) + w_{22}p$:

$$\bar{w} = b_0 + a^2 p (1-p) b_2, \quad (6.40)$$

$$\bar{w}_2 = b_0 - a(1-p)b_1 + a^2(1-p)b_2/2. \quad (6.41)$$

Изменение частоты A_2 за одно поколение задается формулой

$$\Delta p = p(w_2 - \bar{w})/\bar{w}, \quad (6.42)$$

(см. Кроу и Кимура [67], с. 180). Подставляя сюда выражения (6.40) и (6.41) и пренебрегая членами, содержащими a^3 , а также членами более высоких порядков, получим

$$\Delta p = \frac{p(1-p)}{b_0} \left[-ab_1 + a^2 b_2 \left(\frac{1}{2} - p \right) \right]. \quad (6.43)$$

Аналогичное соотношение получено Балмером [32]. Приравняв правые части этого выражения и формулы для изменения частоты гена при генотипическом отборе, т. е. $\Delta p = sp(1-p)$ [см. (6.3) и (6.4)], найдем величину

$$s = -ab_1/b_0 + a^2 b_2 \left(\frac{1}{2} - p \right) / b_0, \quad (6.44)$$

— коэффициент отбора, который отражает преимущество аллеля A_2 перед A_1 . Здесь a — влияние замены A_1 на A_2 на величину признака x .

Применение к стабилизирующему отбору

Стабилизирующий отбор, как известно, является преобладающей формой естественного отбора; к нему мы теперь и обратимся. Предположим, что функция частоты и функции приспособленности задаются нормальным распределением

$$f(x) = \frac{1}{\sqrt{2\pi}} \exp [-(x-m)^2/2], \quad (6.45)$$

$$w(x) = \exp (-kx^2). \quad (6.46)$$

Тогда уравнение (6.44) сводится к

$$s = -m\lambda a + (\lambda^2 m^2 - \lambda) \left(\frac{1}{2} - p \right) a^2, \quad (6.47)$$

где $\lambda = 2k/(1+2k)$. Еще раньше такой же результат получил Балмер [34]. Отметим, что если использовать первоначальный масштаб (X) и представить плотность распределения признака в виде $F(X) = [1/\sqrt{2\pi}\sigma] \exp[-(X-M)^2/(2\sigma^2)]$, функцию приспособленности в виде $W(X) = \exp[-K(X-X_{\text{опт}})^2]$, а эффект замещения аллеля обозначить через A , то параметры в уравнении (6.47) будут равны $m = (M - X_{\text{опт}})/\sigma$, $\lambda = 2K\sigma^2/(1+2K\sigma^2)$ и $a = A/\sigma$.

Из уравнения (6.47) видно, что если $|m| \gg |a|$, т. е. отклонение среднего от оптимума значительно превышает эффект замещения аллеля, то $s/a \approx -\lambda m$, т. е. мы приходим к ситуации, сходной с той, при которой имеет место отсекающий отбор (см. [69, 359]), причем естественный отбор действует весьма эффективно, изменяя среднее в направлении оптимума. В течение этого непродолжительного периода действия направленного отбора может происходить значительный сдвиг частот генов по многим локусам, тогда как генные замены сам этот процесс вызывает редко. С другой стороны, если среднее соответствует оптимуму ($m=0$), то из уравнения (6.47) получаем $s = -\lambda \left(\frac{1}{2} - p \right) a^2$. В этом случае аллель A_2 неблагоприятен, если

$p < 1/2$, и обладает преимуществом, если $p > 1/2$. Такой отбор является частотно-зависимым, а аллели ведут себя как отрицательно сверхдоминантные. Изменение частоты гена в таком случае равно $\Delta p = \lambda a^2 p(1-p)(p-1/2)$ в соответствии с Робертсоном [485] и Райтом [589]. Существенным для нас здесь является тот факт, что каждая новая мутация оказывается вредной, так как $\Delta p > 0$ при малом p . Кроме того, если сегрегирует большое число локусов с малым индивидуальным эффектом, то a становится малой величиной, а a^2 — пренебрежимо малой. Этот эффект выражен особенно сильно, если рассматривать не локусы, а отдельные нуклеотидные сайты, так как было установлено [235], что особь в большой популяции млекопитающих в среднем гетерозиготна, по-видимому, примерно по миллиону нуклеотидных сайтов. Этот результат согласуется с утверждением Оты [435, 437] о том, что большинство мутаций, будучи на молекулярном уровне почти нейтральными, все же оказывают очень слабое вредное действие. Кроме того, можно показать [243], что для отрицательно сверхдоминантных аллелей фиксации в результате случайного дрейфа генов гораздо более вероятна, чем для безусловно вредных с теми же коэффициентами отбора. Интересно и довольно неожиданно, что «балансирующий отбор» на фенотипическом уровне приво-

дит к неравновесному полиморфному состоянию соответствующих генов. В этом смысле утверждения селекционистов по поводу балансирующего отбора на молекулярном уровне в своей основе неверны.

Интенсивность отбора

Мерой интенсивности естественного отбора в стабилизирующей его форме может служить груз L , т. е. доля особей, которые элиминируются в каждом поколении под действием естественного отбора в силу того, что значения фенотипического признака отклоняются у них от оптимума. Для функций частоты и приспособленности, задаваемых формулами (6.45) и (6.46), имеем

$$L = 1 - \sqrt{(1-\lambda) \exp(-\lambda m^2/2)}. \quad (6.48)$$

При $m=0$ это равенство сводится к

$$L = 1 - \sqrt{1-\lambda}. \quad (6.49)$$

В частном случае, когда функция приспособленности имеет ту же дисперсию, что и функция частоты, т. е. когда $K=1/(2\sigma^2)$, или $k=\frac{1}{2}$, получаем $L \approx 0,293$, или примерно 30%-ную элиминацию.

Вообще говоря, у млекопитающих L , по-видимому, должно быть малым для любого отдельного признака. Например, согласно Холдейну [165], интенсивность отбора по весу новорожденных, выражающегося через неонатальную смертность, равна $L=0,027$. Если L мало, то хорошим приближением для (6.49) может служить соотношение

$$L = \lambda/2. \quad (6.50)$$

При $\lambda \leq 0,1$ ошибка такой аппроксимации меньше 2,5%; даже при $\lambda=0,3$ она составляет лишь около 8%. Таким образом, соотношение (6.50) можно использовать для всех практических целей.

Связь общего фенотипического отбора с отбором по индивидуальному локусу

Особи подвергаются естественному отбору по большому числу количественных признаков, многие из которых коррелируют между собой. Чтобы упростить анализ, предположим, что мы можем выбрать некоторое число, скажем $n_{\text{н}}$, независимых признаков, которые все вместе в первом приближении отражают общий характер отбора. Совокупность этих признаков можно назвать полным репрезентативным фенотипом. Различные па-

параметры, относящиеся к i -му признаку, мы будем помечать индексом i ($i=1, 2, \dots, n_{\pi}$).

Так как интенсивность общего отбора ограничена, предполагается, что интенсивность того отбора, который измеряется через L_i и действует на каждый из составляющих признаков, мала, если число соответствующих признаков велико. Пусть L_T — интенсивность общего отбора; тогда

$$(1 - L_T) = \prod_i (1 - L_i), \quad (6.51)$$

так что

$$L_T \approx 1 - \exp\left(-\sum_i L_i\right). \quad (6.52)$$

Чтобы еще больше упростить анализ, предположим, что L_i одинаков для всех признаков; тогда $L_i \approx -(1/n_{\pi}) \cdot \ln(1 - L_T)$. Коэффициент отбора в расчете на локус (или сайт) равен $-\lambda_i(1/2 - p_i)a_i^2$, где $\lambda_i = 2k_i/(1 + 2k_i)$. Учитывая, что $L_i \approx \lambda_i/2$ и опуская индекс, получаем

$$s \simeq [\ln(1 - L_T)](1 - 2p)a^2/n_{\pi}. \quad (6.53)$$

Заметим, что в правой части этого выражения $a^2/n_{\pi} = A^2/(n_{\pi}\sigma^2)$, где $n_{\pi}\sigma^2$ — дисперсия фенотипа в целом.

Интересно было бы оценить, насколько велик отбор в расчете на нуклеотидный сайт, если все сайты вносят одинаковый вклад в фенотип. Пусть $\bar{h}$ — средняя гетерозиготность на сайт; если мы обозначим полное число участвующих сайтов через $n_{\text{нукл}}$, то $n_{\text{нукл}}A^2\bar{h} = n_{\pi}\sigma^2\rho^2$, где A — эффект замены одного нуклеотида на определенный компонент сложного фенотипа, а ρ^2 — наследуемость в широком смысле. Таким образом, коэффициент стабилизирующего отбора s_c , определяемый из выражения $s = -s_c(1 - 2p)$, можно представить следующим образом:

$$s_c = -[\ln(1 - L_T)]\rho^2/(n_{\text{нукл}}\bar{h}). \quad (6.54)$$

Это отражает интенсивность стабилизирующего отбора при нуклеотидном замещении (если $m=0$). Чтобы оценить s_c , предположим, что средняя гетерозиготность на ферментный локус, относящаяся к электрофоретически определяемым аллелям, равна 0,1, а гетерозиготность по молчащим аллелям вдвое выше. Тогда, если среднее число нуклеотидов, из которых состоит данный локус, равно 10^3 , получим $\bar{h} = 3 \cdot 10^{-4}$. Экстраполируя эти данные на весь геном млекопитающих, состоящий из $3,5 \cdot 10^9$ нуклеотидов, получим, что среднее число гетерозиготных нуклеотидных сайтов на особь составляет $n_{\text{нукл}}\bar{h} = 1,05 \cdot 10^6$. Интенсивность общего отбора для крупных млекопитающих не может быть очень высокой. Диви [79] рассчитал, что у челове-

ка на протяжении всего периода его эволюции сложился такой тип воспроизводства, при котором около половины всех рождавшихся детей доживало до половой зрелости. Таким образом, типичный груз для млекопитающих (например, для человека) равен, по-видимому, $L_T=0,5$. Что касается наследуемости в широком смысле, то она предположительно равна $\rho^2=0,5$. Таким образом, получаем $s_c=3,3 \cdot 10^{-7}$, т. е. коэффициент стабилизирующего отбора крайне мал, и большинство мутаций на молекулярном уровне почти нейтральны, хотя и обладают очень слабым вредным эффектом. Это дает большую свободу действия случайному генетическому дрейфу. Полученные выводы согласуются с гипотезой Оты [435, 437] об очень слабом вредном эффекте мутаций (однако приспособленность вида не уменьшается, как это должно было быть в соответствии с этой гипотезой). При этом считается, что гены, которые замещаются в результате случайного дрейфа, и гены, ответственные за фенотипическую изменчивость количественных признаков, принадлежат к одному классу. Вероятно, многие (если не большинство) мутации, влияющие на данный количественный признак, затрагивают регуляторные, а не структурные гены. С этой точки зрения некодирующие участки ДНК могут быть более важными, чем транслируемые. Приведенные рассуждения согласуются с мнением Ланде [290], который считает, что в результате случайного дрейфа может накапливаться много полигенных изменений, так как их фенотипический эффект либо мал, либо вообще отсутствует. Излишне говорить, что одни сайты могут приводить к гораздо большим фенотипическим эффектам, чем другие, и поэтому подвергаются более сильному отбору. С другой стороны, определенная (и, по-видимому, большая) доля сайтов вообще не обладает фенотипическими эффектами и, следовательно, совершенно нейтральна по отношению к естественному отбору.

6.8. Поведение мутантных аллелей в конечной популяции при стабилизирующем отборе

Вероятность фиксации мутантного аллеля при стабилизирующем отборе

Прежде чем перейти к проблемам селективного ограничения и скорости эволюции, рассматриваемым в следующей главе, имеет смысл оценить последствия случайного дрейфа при стабилизирующем отборе — основном типе естественного отбора. Это, в частности, поможет нам понять причину «неслучайного использования кодонов» (разд. 7.5) в рамках теории нейтральной генетики, основанной на стохастической теории популяционной генетики.

В разд. 3.5 мы выяснили, что скорость эволюции, выражен-

ная через число мутационных замен, задается выражением $k = 2Nvu$ [см. соотношение (3.19)], где N — численность популяции, v — скорость мутирования, u — вероятность фиксации отдельной мутации. В разд. 6.7 этой главы показано, что в условиях стабилизирующего отбора каждая мутация обладает лишь небольшим вредным эффектом. Таким образом, чтобы оценить эффективность случайного дрейфа при стабилизирующем отборе, необходимо найти вероятность u фиксации мутантного аллеля в этих условиях.

Рассмотрим некий количественный признак, определяемый большим числом локусов, каждый из которых имеет малый эффект. Предположим, что в диплоидной популяции со случайным скрещиванием по данному локусу сегрегирует пара аллелей A_1 и A_2 с относительными частотами $1-p$ и p . Предположим далее, что A_2 — мутантный аллель, образовавшийся из A_1 . Обозначим через N число особей, составляющих популяцию, и через N_e — эффективную численность (о смысле N_e см. разд. 3.3; вообще говоря, N_e — это число особей, размножающихся в одном поколении). В большинстве случаев N_e , по-видимому, много меньше общей численности популяции N .

Как показано в предыдущем разделе, если функции частоты и приспособленности следуют нормальному распределению [см. (6.45) и (6.46)] и если a — эффект замены A_1 на A_2 , то согласно (6.47),

$$s = \beta_1 - \beta_2(1 - 2p), \quad (6.55)$$

где $\beta_1 = -\lambda m a$, $\beta_2 = \lambda(1 - \lambda m^2)a^2/2$, $\lambda = 2k/(1 + 2k)$.

Пренебрегая пока давлением мутаций, обозначим частоту A_2 через x (не путайте с обозначениями в предыдущем разделе, где x использовалось для представления величины признака в единицах σ). Тогда среднее и дисперсия изменения x за одно поколение равны

$$M_{\delta x} = [\beta_1 - \beta_2(1 - 2x)]x(1 - x), \quad (6.56)$$

$$V_{\delta x} = x(1 - x)/(2N_e). \quad (6.57)$$

Пусть $u(p)$ — вероятность фиксации аллеля A_2 , начальная частота которого составляла p . Тогда $u(p)$ можно выразить через $M_{\delta x}$ и $V_{\delta x}$, используя основную формулу вероятности фиксации, полученную Кимурой [227]:

$$u(p) = \int_0^p G(x) dx \bigg/ \int_0^1 G(x) dx, \quad (6.58)$$

где

$$G(x) = \exp \left\{ -2 \int \frac{M_{\delta x}}{V_{\delta x}} dx \right\} \quad (6.59)$$

(см. [67], с. 424, или [257], с. 173).

Особый интерес представляет вероятность фиксации аллеля A_2 , который вначале был представлен в популяции в единственном числе. Эта вероятность, обозначаемая через u , задается функцией $u(p)$, где $p = 1/2N$.

$$u = \frac{1}{2N \int_0^1 \exp[-B_1x + B_2x(1-x)] dx}, \quad (6.60)$$

где $B_1 = 4N_e\beta_1$, $B_2 = 4N_e\beta_2$.

В приведенных выше выкладках мы предполагали, что m , т. е. отклонение среднего от оптимума [см. рис. 6.6; заметим, что $m = (M - X_{\text{опт}})/\sigma$], остается все время неизменным. Такое предположение, по-видимому, нереалистично, так как если $m \neq 0$, то $|m|$ должно уменьшаться под действием направленного компонента отбора гораздо быстрее, чем мутантный ген дрейфует к состоянию фиксации. Этому изменению, однако, может противостоять давление мутаций; в результате m в ходе стабилизирующего отбора останется постоянным, хотя в равновесии $|m|$, по-видимому, мало. Такая ситуация наблюдается в том случае, когда оптимум и точка мутационного равновесия не совпадают. Мы встретимся с ней в следующей главе, когда будем обсуждать проблему «неслучайного» использования синонимичных кодонов.

Мы уже видели, что отбор, действующий на мутантный аллель, является частотно-зависимым и что аллели ведут себя как отрицательно сверхдоминантные. Это означает, что мутации, испытывающие действие стабилизирующего отбора, должны фиксироваться путем случайного дрейфа с гораздо большей вероятностью, чем вредные мутации, имеющие сравнимые коэффициенты отбора. Чтобы показать, что это действительно так, я привожу в табл. 6.3 несколько примеров вероятностей фиксации для этих двух случаев. Для стабилизирующего от-

Таблица 6.3. Относительная вероятность фиксации (u/u_0) мутаций с отрицательным сверхдоминированием и вредных мутаций (u_0 — вероятность фиксации полностью нейтральной мутации). $\bar{S}$ равно $4N_e s_c$ для мутаций с отрицательным сверхдоминированием и $4N_e s'$ — для вредных мутаций

S	Мутации с отрицательным сверхдоминированием	Вредные мутации
0	1,00	1,00
1,0	0,84	0,58
8,0	0,23	$2,7 \cdot 10^{-3}$
16,0	0,042	$1,8 \cdot 10^{-6}$
20,0	0,017	$4,1 \cdot 10^{-8}$
30,0	0,0017	$2,8 \cdot 10^{-12}$

бора мы принимаем $m=0$ и обозначаем $\lambda a^2/2$ через s_c (т. е. через коэффициент стабилизирующего отбора), так что $\Delta p = -s_c p(1-p)(1-2p)$. Для вредной мутации мы обозначаем коэффициент отбора против A_2 через $-s'$ ($s' > 0$), так что вероятность фиксации задается выражением

$$u = S' / [2N(e^{S'} - 1)], \quad (6.61)$$

где $S' = 4N_e s'$ (см. [67], с. 426). В табл. 6.3 приведены значения u/u_0 для обоих случаев (где $u_0 = 1/2N$ — вероятность фиксации нейтрального аллеля). Видно, что вероятности фиксации в этих двух случаях совершенно различны и что при стабилизирующем отборе может иметь место нейтральная эволюция, даже если $4N_e s_c \geq 8$. Можно показать, что при $B = 4N_e s_c > 8$ $u/u_0 \approx \sqrt{(B/\pi) \exp(-B/4)}$.

Распределение частот генов

Учтем теперь давление мутаций и исследуем вероятностное распределение аллельных частот при статистическом равновесии, достигаемом конечной популяцией под действием стабилизирующего отбора. Обозначим через $\phi(x)$ плотность вероятности, так что $\phi(x)dx$ будет представлять собой вероятность того, что частота A_2 в популяции находится в интервале от x до $x+dx$, где $0 < x < 1$. Обозначим через v_1 скорость мутирования A_1 к A_2 , а через v_2 — скорость обратного мутирования. Тогда среднее и дисперсия изменения частоты A_2 в одном поколении будут равны

$$M_{\delta x} = [\beta_1 - \beta_2(1-2x)]x(1-x) - v_2 x + v_1(1-x) \quad (6.62)$$

и

$$v_{\delta x} = x(1-x)/(2N_e). \quad (6.63)$$

Воспользуемся формулой Райта [591, 592] для стационарного распределения частот генов:

$$\phi(x) = \frac{C}{V_{\delta x}} \exp \left\{ 2 \int \frac{M_{\delta x}}{V_{\delta x}} dx \right\}, \quad (6.64)$$

где константа C определяется так, что

$$\int_0^1 \phi(x) dx = 1 \quad (6.64a)$$

(см. [67], с. 434—435). Отсюда получаем

$$\phi(x) = C e^{B_1 x - B_2 x(1-x)} \cdot x^{V_1-1} \cdot (1-x)^{V_2-1}, \quad (6.65)$$

где $B_1 = 4N_e \beta_1$, $B_2 = 4N_e \beta_2$, $V_1 = 4N_e v_1$, $V_2 = 4N_e v_2$.

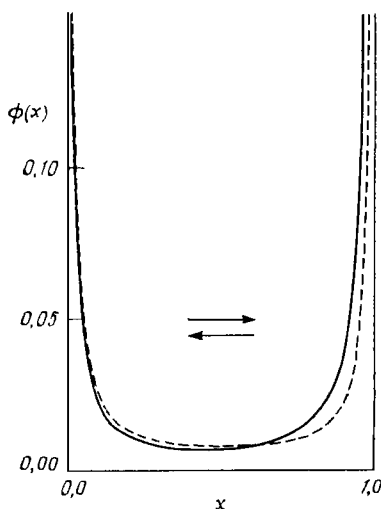


Рис. 6.8. Распределение плотности вероятности (ϕ) частоты (x) аллеля A_2 . Сплошная кривая — распределение при стабилизирующем отборе с $V_1=V_2=4 \cdot 10^{-3}$, $B_1=0,847$, $B_2=2$ [см. формулу (6.65)]. Это распределение относится только к тем случаям, когда в популяции присутствуют оба аллеля ($0 < x < 1$). В действительности в подавляющем большинстве случаев при тех значениях параметров, которые приведены выше, аллель A_2 либо фиксирован в популяции ($x=1$), либо утрачен ($x=0$). Симметричная штриховая линия изображает полностью нейтральную ситуацию ($B_1=B_2=0$) при тех же скорости мутирования и численности популяции ($V_1=V_2=4 \cdot 10^{-3}$). Горизонтальные стрелки показывают, что в ходе длительной эволюции происходит переход от одного крайнего состояния (когда, скажем, фиксирован A_1) к другому (когда фиксируется A_2) через состояние полиморфизма.

Тогда вероятность того, что аллель A_2 на какое-то время зафиксирован в популяции, можно найти, проинтегрировав $\phi(x)$ от $1-1/2N$ до 1; таким образом, с достаточной точностью получаем

$$f_2 = C e^{B_1/[V_2(2N)^{V_2}]} \quad (6.66)$$

Аналогичным образом вероятность того, что аллель A_1 на какое-то время зафиксирован в популяции (т. е. A_2 будет утрачен), равна

$$f_1 = C/[V_1(2N)^{V_1}] \quad (6.67)$$

Отношение f_2 к f_1 будет равно

$$f_2/f_1 = e^{B_1(V_1/V_2)(2N)^{V_1-V_2}} \quad (6.68)$$

В этой книге мы будем говорить в основном о таких ситуациях, когда и V_1 , и V_2 много меньше единицы, а аллели большую часть времени фиксированы. Именно такая картина наблюдается, когда рассматривается отдельный нуклеотидный

сайт, а не обычный генный локус, поскольку скорость мутирования на сайт должна быть порядка 10^{-8} , а не 10^{-5} .

В качестве примера рассмотрим нуклеотид, находящийся в третьем положении кодона. Сгруппируем четыре основания в два класса, обозначив основания А и U через A_1 , а С и G — через A_2 . В этом сайте могут находиться два «аллеля», и распределение частоты A_2 задается уравнением (6.65). Предположим, что скорость мутирования в обоих направлениях одинакова и равна $\nu_1 = \nu_2 = 10^{-8}$ на поколение, а эффективная численность популяции составляет 10^5 , так что $V_1 = V_2 = 4 \cdot 10^{-3}$. При таких условиях аллели находятся преимущественно в фиксированном состоянии. Допустим далее, что фиксация A_2 происходит чаще, чем A_1 , и отношение $f_2/f_1 \approx 7/3$, как в генах гемоглобина. Тогда из равенства (6.68) находим $B_1 = \ln(7/3) = 0,847$. Что касается эффекта единичной мутационной замены, то будем считать его очень малым: $a = 10^{-2}$, так что $a^2 = 10^{-4}$. Это соответствует $3 \cdot 10^4$ генам, кодирующим белки, со средней гетерозиготностью 0,17 на локус и наследуемостью в широком смысле (ρ^2) 50%. По отношению к интенсивности стабилизирующего отбора на оптимизацию использования кодонов (разд. 7.5) интенсивность изменения в третьем положении кодона должна быть нулевой или очень малой. Поэтому можно допустить, что $L = 0,05$ (1/10 общего груза), так что $\lambda = 0,1$. Заметив, что $\lambda m^2 \ll 1$, получаем $B_2 = 4N_e \beta_2 \approx 4N_e \lambda a^2 / 2 = 2$. На рис. 6.8 приведено распределение частот аллеля A_2 .

Структура молекул и скорость эволюции

7.1. Консервативная природа мутационных замен

Уже из ранней работы Цукеркандла и Полинга, опубликованной в 1965 г. [622], стало известно, что в ходе эволюции белков чаще всего происходят взаимные замены аминокислот, лишь незначительно различающихся по структуре и химическим свойствам. По мнению авторов, такие «консервативные» замены, вероятно, сопровождаются весьма небольшими изменениями функциональных свойств белковых молекул и поэтому должны «удовлетворять естественному отбору». Однако они же отмечали, что, по-видимому, «химики и белковые молекулы имеют разные мнения» относительно того, какие из свойств аминокислотных остатков наиболее существенны. Было высказано предположение, что лучшим свидетельством консервативности замены является ее высокая частота в эволюции белков. У такого рассуждения есть один недостаток: оно легко приводит к замкнутому кругу. Необходимо найти более объективную меру сходства аминокислот и затем исследовать, существует ли между таким сходством и частотой замен в ходе эволюции положительная корреляция. Другой момент, на который следует обратить внимание в такого рода анализе, состоит в том, что генетический код обладает одним интересным свойством, отмеченным еще Цукеркандлом и Полингом [622]: замена одного основания в кодоне часто приводит к сходной аминокислоте или к синонимичному кодону. Другими словами, код сам по себе «предрасположен» к консервативным мутационным изменениям. По-видимому, первым, кто исследовал консервативную природу замен, используя объективный показатель для представления различий между парами аминокислот, был Эпштейн [95]. Этот показатель (названный коэффициентом различия) основан на различиях в полярности и размере аминокислот; автор сравнил его распределение для аминокислотных замен в гомологичных белках (субилизине, α - и β -цепях гемоглобина, цитохроме *c*, рибонуклеазе и др.). Результаты исследования показали, что для всех проведенных сравнений гомологичных белков коэффициенты различия были значительно ниже, чем

можно было ожидать для «случайных мутаций», даже если принять во внимание консервативную природу генетического кода.

Позднее Кларк [49] провел подобное исследование, используя коэффициент Снита D [520], отражающий химическое различие между аминокислотами, и данные Дэйхоф по относительным частотам аминокислотных замен [178]. Учтя различия в частотах и в вероятностях мутационных изменений (т. е. различия в числе синонимичных кодонов для разных аминокислот), он обнаружил статистически значимую отрицательную корреляцию между частотой замен (R) и химическим различием (D). Это показывает, что аминокислотные замены, сопровождающиеся небольшими изменениями химических свойств, происходят в ходе эволюции гораздо чаще, чем замены, приводящие к существенным их изменениям. Кларк обнаружил также, что средняя скорость замен составляет лишь около $1/9$ от максимальной скорости, которая наблюдается при $D=0$. Он заявил, что эти наблюдения говорят о несостоятельности ряда «менее критических» исследований по эволюции белков (в число которых он включил работы Кинга и Джукса [268] и, вероятно, Кимуры [229, 231]). Согласно Кларку, 85—90% мутаций, приводящих к заменам аминокислот, являются вредными, но эта цифра учитывает только один вид селективного ограничения. В действительности же, замечает он, имеется немало других, таких, например, как ограничения, налагаемые типом белка, положением данной аминокислоты в белке, видом организма, которому принадлежит этот белок. Если учесть все эти факторы, то окажется, что средняя доля «нейтральных» мутаций приближается уже к 0,1%, а не к 10%. Джукс и Кинг [208], опровергая доводы Кларка, подчеркивали, что если селективно нейтральные и почти нейтральные изменения обычны для эволюции, то замены на химически сходные аминокислоты должны происходить гораздо чаще, чем на несходные (что и соответствует наблюдениям). С другой стороны, чтобы могла осуществляться положительная дарвиновская эволюция, модифицированный белок должен отличаться функционально от исходного и радикальные изменения должны быть более благоприятными, чем консервативные. В свою очередь Кларк [50], ссылаясь на Фишера [107], отметил, что малое мутационное изменение имеет больше шансов оказаться благоприятным (см. разд. 6.6) и что отбор скорее всего работает не в пользу резких изменений. Мне кажется, Кларк не учел того обстоятельства, что мутации с очень малым эффектом, evidentemente, имеют весьма небольшое преимущество при отборе, а потому и небольшую вероятность фиксации, т. е. у них мало шансов распространиться по популяции данного вида. Как мы покажем позже, это означает, что аргументы Фишера нельзя использовать для объяснения с позиций селекционизма того

факта, что, чем меньше различие между аминокислотами, тем выше скорость их замен в ходе эволюции.

В рамках теории нейтральности природу консервативных замен легко объяснить, приняв во внимание, что, чем меньше различие между двумя аминокислотами, тем выше вероятность селективной эквивалентности соответствующих мутаций, а отнюдь не вредности. (Теория нейтральности в ее крайней форме предполагает, что ни одна из мутаций не имеет преимуществ перед другой.) Следовательно, вероятность взаимозамены с селективно нейтральными последствиями выше для сходных аминокислот, а значит, такие замены, осуществляющиеся в ходе эволюции благодаря случайному генетическому дрейфу, должны происходить чаще. Пользуясь обозначениями, принятыми в уравнении (5.1), можно сказать, что доля f_0 больше для мутаций, приводящих к заменам на сходные аминокислоты.

Иногда теорию нейтральности трактуют неверно, полагая, что замены разных типов аминокислот должны происходить с одинаковой частотой (для чего обычно применяют термин «случайность»), наблюдаемое же неравенство частот («неслучайность») рассматривается как свидетельство против теории нейтральности. Грэнсем [143], например, утверждает, что противоречие, касающееся механизма молекулярной эволюции, — это по существу противоречие между «случайностью и физико-химическим детерминизмом», и отождествляет сторонника нейтралистской теории с «защитником случайности». Для того чтобы исследовать консервативную природу аминокислотных замен, Грэнсем [143] ввел формулу, отражающую физико-химические различия между аминокислотами, аналогично тому, как это было сделано Эпштейном [95] и Кларком [49]. В основе его формулы для D_{ij} — меры «расстояния» между двумя аминокислотами i и j — лежат три параметра: полярность аминокислот, их объем и состав, причем состав определяется как отношение атомных масс неуглеродных атомов, входящих в концевые группы или кольца, к атомным массам углеродов боковой цепи. Он исследовал также корреляцию между D_{ij} и относительной частотой замен (RSF) (см. табл. A1 в дополнении к работе [323]). Оказалось, что линейная регрессия $\lg \text{RSF}$ на D имеет коэффициент корреляции $-0,72$, что означает гораздо более сильную отрицательную корреляцию, чем та, которая была получена с использованием коэффициента Снита и формулы Эпштейна. Однако эти данные согласуются с результатами предыдущих работ в том отношении, что в соответствии с ними фиксация несинонимичных мутаций происходит, как правило, редко.

Позже Мията и др. [362] предложили простую меру «расстояния» между аминокислотами, основанную только на двух параметрах — полярности (p) и объеме (v), каждый из кото-

рых был взвешен по величине, обратной их стандартным отклонениям σ_p и σ_v . Согласно этим авторам, если рассматривать только те аминокислотные замены, которые можно отнести к одноступенчатым мутациям, то для меры их расстояния будет наблюдаться сильная отрицательная корреляция с наблюдаемой частотой замен, что ясно указывает на консервативную природу замен в ходе эволюции. Анализ данных по аномальным гемоглобинам человека, проведенный этими авторами, выявил положительную корреляцию между частотой замен и степенью различий аминокислот. Это и понятно, так как в случае аномальных гемоглобинов легче всего регистрируются и чаще всего описываются именно те аминокислотные замены, которые приводят к наиболее серьезным патологическим эффектам.

Консервативную природу аминокислотных замен в случае миоглобинов млекопитающих исследовали также Богард и др. [25]. Используя структурную дискриминантную функцию D_s , зависящую от четырех параметров — объема, полярности, изоэлектрической точки и гидрофобности аминокислот, — и проводя сравнение последовательностей известных миоглобинов млекопитающих, они установили, что в ходе эволюции, по-видимому, происходили в основном такие аминокислотные замены, которые не сопровождались изменением исходной конформации белковой молекулы. В частности, ими показано, что коэффициент Δ_m , который по существу является средним величины D_s для аминокислотного сайта, больше для «внешних», чем для «внутренних» аминокислотных остатков, а это означает, что в случае внешних остатков возможны менее консервативные (т. е. более радикальные) замены. Это легко понять, поскольку для менее важной в функциональном отношении внешней части белковой молекулы могут оказываться приемлемыми (т. е. селективно нейтральными или эквивалентными) более разнообразные аминокислоты, и это не приведет к нарушению функциональных свойств молекулы. Однако авторы считают, что большинство «нейтральных» аминокислотных замен есть результат «селективной фиксации приемлемых альтернативных стабилизирующих мутаций». Как было показано в предыдущей главе (разд. 6.8), при стабилизирующем отборе может протекать достаточно эффективная нейтральная эволюция, но происшедшие при этом мутационные замены нельзя назвать «селективной фиксацией», поскольку фиксация мутаций обусловливается генетическим дрейфом, а не отбором. Фактически действующий при этом отрицательный отбор крайне слаб, и случайный дрейф генов может легко перекрыть его.

Чтобы проиллюстрировать более убедительно консервативную природу аминокислотных замен с позиций популяционной генетики, я построил график зависимости физико-химических различий между аминокислотами от относительной частоты

их замен (рис. 7.1). На этом рисунке по оси абсцисс отложено расстояние, введенное Миятой и др. [362], а по оси ординат — относительная частота аминокислотных замен (табл. А1 в работе [323]). Данные получены в результате сравнения 17 семейств гомологичных белков, в том числе 13 гемоглобинов, 8 цитохромов *c*, 14 варьибельных участков легких цепей

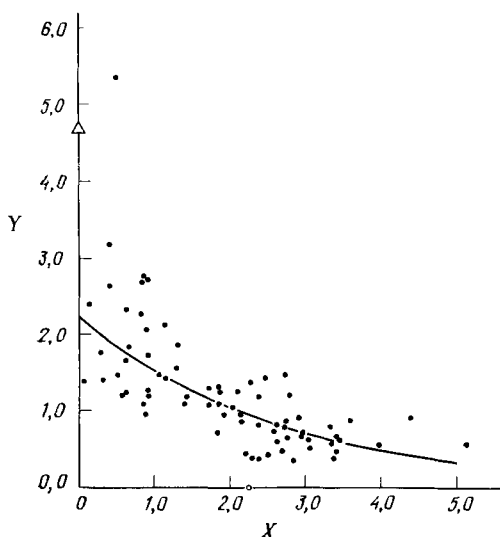


Рис. 7.1. Зависимость между физико-химическими различиями (X) и относительной частотой замен (Y) аминокислот при эволюции белков. Точки — наблюдаемые частоты, кривая — экспоненциальная функция $Y = 2,22e^{-0,376X}$. Треугольник на оси Y отвечает ожидаемой частоте синонимичных замен, рассчитанной по данным, полученным при сравнении нуклеотидных последовательностей генов гемоглобинов (подробности см. в тексте).

антител и т. д. По-видимому, это наиболее полная на сегодняшний день сводка данных по частотам аминокислотных замен. Заметим, что этот показатель нормирован с учетом того, что разные аминокислоты встречаются с разными частотами. Наблюдаемые частоты представлены точками; сплошная кривая — построенная по этим точкам экспоненциальная функция $Y(X) = Ae^{BX}$. Из 20 аминокислот можно образовать 190 комбинаций (аминокислотных пар), но я выбрал только 75 пар, которые связаны одним мутационным шагом, т. е. те пары аминокислот, которые переходят одна в другую при замене лишь одного основания в кодоне. Кроме того, подгоняя к этим данным экспоненциальную функцию $Y = Ae^{BX}$ (или, что то же самое, линейную функцию $\ln Y = \ln A + BX$), я исключил одну пару (Cys $\leftrightarrow$ Phe, светлый кружок на рис. 7.1), для которой нет

данных в таблице A1 работы [323]. Линейная регрессия $\ln Y$ на X в свою очередь дает наклон прямой $B = -0,376 \pm 0,04$ и ординату точки пересечения с осью $Y \ln A = 0,797 \pm 0,09$. Коэффициент корреляции равен $-0,742$, а среднее значение $\bar{X} = 1,93$. Относительная частота, отвечающая среднему расстоянию между членами пары аминокислот, составляет $Y(\bar{X}) = 1,07$; это меньше чем половина $Y(0) = 2,22$.

С позиций теории нейтральности особенно интересно выяснить, соответствует ли относительная частота, полученная путем экстраполяции к $X=0$, той частоте, которая ожидается исходя из синонимичных замен (т. е. замен, которые не приводят к замене аминокислот). С этой целью оценим вначале отношение $R_{S/A}$ числа синонимичных замен к несинонимичным, приходящимся на сайт и на эквивалентный выход мутаций, используя данные, полученные при сравнительном анализе нуклеотидных последовательностей. Как уже указывалось в последнем подразделе гл. 4, расчетное число аминокислотных замен, обусловленных синонимичными изменениями в третьем положении кодона, задается величиной K'_s . Заметим, что синонимичными по третьему положению кодона являются примерно $(1 + \frac{1}{3})/2$, или $\frac{2}{3}$, случайных нуклеотидных замен, так как для любого данного сочетания оснований в первом и втором положениях кодона изменение третьего основания примерно в половине случаев полностью синонимично, а в остальной половине случаев лишь частично синонимично (пурин $\longleftrightarrow$ пурин или пиримидин $\longleftrightarrow$ пиримидин). С другой стороны, замещение нуклеотида в первом и втором положениях кодона почти всегда приводит к замене аминокислоты. Таким образом, ожидаемую величину $R_{S/A}$ можно оценить как отношение $K'_s / (\frac{2}{3})$ к $(K_1 + K_2)/2$, т. е. $3K'_s / (K_1 + K_2)$. Воспользуемся данными табл. 4.8 для гемоглобинов (где представлены стандартные скорости аминокислотных замен), не рассматривая при этом сопоставления, включающие аномальный α -3-ген («молчащий» ген), и сопоставления α - и β -цепей (которые различаются слишком сильно, чтобы можно было точно определить K'_s). Тогда величина $3K'_s / (K_1 + K_2)$ оказывается равной $4,43 \pm 1,64$. Если мы умножим ее на $Y(\bar{X})$, то получим грубую оценку частоты синонимичных замен. Эта частота равна $4,7 \pm 1,8$ и отмечена светлым треугольником на рис. 7.1. Хотя она значительно превышает величину $Y(0) = 2,22$, что говорит о большой статистической ошибке такого метода расчета, все же можно сделать вывод, что данные по синонимичным заменам согласуются с данными по белкам о зависимости между скоростью эволюции и масштабом изменений, т. е., чем меньше мутационное изменение, тем выше скорость эволюции, выражаемая через число мутационных замен.

Это положение легко объяснить в рамках теории нейтральности, используя уравнение (5.1), т. е. $k = v_T f_0$. Вероятность f_0 селективно нейтрального мутационного изменения становится больше, если соответствующая замена аминокислоты связана с меньшим изменением физико-химических свойств, и величина k соответственно увеличивается.

Посмотрим теперь, какова должна быть зависимость между k и x (где x — величина мутационного изменения), если предположить, что мутационные замены обусловлены исключительно положительным дарвиновским отбором. При таком допущении k пропорционально произведению $P_a(x)u_a(x)$, где $P_a(x)$ — вероятность того, что мутация с эффектом x окажется селективно выгодной, а $u_a(x)$ — вероятность появления такой мутации, которая в конце концов зафиксировается в популяции.

Если предположить, что селективное преимущество s пропорционально x , тогда $u_a(x) \propto 2x$, поскольку в соответствии с уравнением (3.21) $u \approx 2sN_e/N$. Что касается $P_a(x)$, то мы вслед за Фишером [107] положим, что она задается формулой (6.33), т. е. $P_a(x)$ уменьшается с увеличением x (см. рис. 6.5). Следовательно,

$$k \propto 2xP_a(x) \quad (7.1)$$

(см. рис. 7.2). В отличие от данных, представленных на рис. 7.1, эта кривая имеет колоколообразную форму, причем максимум скорости эволюции достигается при промежуточном значении мутационного эффекта, что противоречит утверждению Кларка [52]. Отметим также, что, хотя мутации с меньшими эффектами имеют больше шансов стать благоприятными (см. работу Фишера [107]), их селективное преимущество невелико. В случае обычных фенотипических признаков (количественных признаков) необходимо учесть еще один фактор, а именно: мутации с малыми фенотипическими эффектами могут встречаться чаще, чем с большими. Рассматривая эволюцию на молекулярном уровне, последнее можно исключить, как мы уже сделали выше, учитывая различие в частотах аминокислот и/или различие в числе синонимичных кодонов для разных аминокислот.

Можно было бы возразить, что селективное преимущество s не становится пренебрежимо малым, даже если мутационный эффект x уменьшается до нуля; напротив, оно стремится к некой положительной величине s_0 , гораздо большей, чем $1/2 N_e$. В таком случае $P_a(x)$ должно стремиться к $1/2$ при $x \rightarrow 0$, т. е. синонимичные мутации имеют селективное преимущество в 50% случаев. Тогда $f_A = 0,5$, $s_A = s_0$, и для синонимичных замен мы имеем $k = 2N_e s_0 v_T$. Следовательно, k гораздо больше полной скорости благоприятных мутаций v_T . Этот вывод полностью противоположен предсказанию теории нейтральности, согласно которому k меньше чем v_T , хотя и может приближаться к не-

му. Однако в разд. 7.3 мы покажем, что для синонимичных замен характерно одно примечательное свойство: их частота не только высока, но и примерно одинакова для разных генов, даже когда частоты замен, изменяющих аминокислоты, сильно

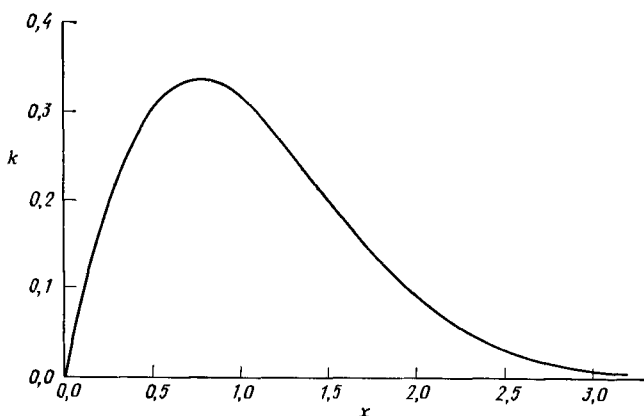


Рис. 7.2. Зависимость между величиной мутационного изменения x (в произвольных единицах) и относительной скоростью эволюции (выраженной через число мутационных замен) для случая, когда мутационные замены обусловлены исключительно дарвиновским отбором.

различаются; это находится в значительно большем соответствии с теорией нейтральности, чем выводы селекционистской теории.

7.2. Молекулярная структура и селективные ограничения

Становится все более очевидным, что для молекулярной эволюции характерна не только консервативная природа мутационных замен, но и более высокая частота изменений менее важных в функциональном отношении молекул или их участков по сравнению с более важными. Среди исследованных белков наибольшая скорость эволюции наблюдается у фибринопептидов (см. табл. 4.1). По оценкам, она составляет $8,3 \cdot 10^{-9}$ на аминокислотный сайт в год; правда, в работе [18] приводится меньшая величина — $4,5 \cdot 10^{-9}$. С другой стороны, самая низкая скорость эволюции характерна для гистона H4 — $0,008 \cdot 10^{-9}$ на аминокислотный сайт в год. Эту скорость можно непосредственно вычислить, исходя из того хорошо известного факта, что гистон H4 гороха (состоящий примерно из 100 аминокислот) отличается от гистона тимуса телянка только двумя аминокислотами [193], и предположив, что растения и животные дивергировали около 1,2 млрд. лет назад. Интересно отметить, что та-

кие быстро эволюционирующие молекулы, как молекулы фибринопептида, после своего отщепления от фибриногена в процессе образования кровяного сгустка, насколько нам известно, по-видимому, не выполняют каких-либо существенных функций. Тем не менее они не являются полностью инертными и обладают известным физиологическим действием. Однако функционально важен лишь небольшой участок полипептидной цепи [288]; в остальной же ее части могут происходить разнообразные изменения. К этой проблеме мы еще вернемся несколько позже.

Связь между функциональным значением (или, более строго, функциональным ограничением) и скоростью эволюции прекрасно объяснил Дикерсон [80]. У фибринопептидов практически любая мутационная замена аминокислоты, которая не препятствует их отщеплению, «приемлема» для вида. Таким образом, скорость эволюционных замен аминокислот может быть очень близкой здесь к истинной скорости мутирования. Гемоглобины выполняют вполне определенную функцию, осуществляя перенос кислорода, поэтому требования к их структуре более строгие, чем у фибринопептидов, и скорость эволюции ниже. Цитохром *c* при своем функционировании взаимодействует с цитохромоксидазой и редуктазой — гораздо более крупными молекулами, — и функциональные ограничения здесь еще более жесткие, чем у гемоглобинов, а соответственно и скорость эволюции меньше. Гистон H4 находится в комплексе с ядерной ДНК, и совершенно очевидно, что белок, столь тесно связанный с молекулой — носителем генетической информации, должен быть высоко специфичным и подвергаться очень небольшим изменениям за миллионы лет эволюции. С точки зрения теории нейтральности эти положения можно объяснить следующим образом. Вероятность нейтральных мутационных изменений жестко связана с молекулярными ограничениями. Если молекула или ее часть функционально менее значима, то вероятность появления селективно нейтрального (т. е. селективно эквивалентного) изменения будет выше. Иными словами, одна и та же функция может выполняться одинаково хорошо при участии разных аминокислот. Таким образом, скорость эволюции на уровне мутационных замен благодаря роли случайного дрейфа оказывается высокой.

Кроме того, чем меньше данная молекула подвержена действию естественного отбора, тем выше вероятность того, что возникшая в ней мутация будет селективно нейтральной. С этим согласуется тот факт, что для δ -цепи гемоглобина $A_2(\alpha_2\delta_2)$, минорного компонента гемоглобина взрослого человека, характерны более высокая скорость эволюции и более высокий уровень полиморфизма, чем для β -цепи главного компонента гемоглобина — гемоглобина A ($\alpha_2\beta_2$) [27].

Скорость эволюции может быть неодинаковой не только для разных молекул, но и для разных участков одной и той же молекулы. Рассмотрим, например, гемоглобин — один из белков, чья структура хорошо известна. В молекуле этого белка имеется несколько спирализованных и неспирализованных участков; как и в других глобулярных белках, гидрофобные аминокислоты находятся внутри структуры и мало изменяются в ходе эволюции. Ключевую роль в функционировании гемоглобина играет та часть молекулы, которая примыкает к гему (так называемый гемовый «карман»), и происходящие здесь замещения аминокислот обычно приводят к аномальным изменениям свойств гемоглобина, например к нестабильности [467]. С другой стороны, аминокислоты, располагающиеся на поверхности молекулы, обычно не играют особой функциональной роли, и замены мутантными аминокислотами в этой области часто не дают никаких клинических симптомов в гетерозиготном состоянии.

Чтобы сравнить скорости эволюции для этих двух участков [259], мы проанализировали данные об аминокислотной последовательности α - и β -цепей (в основном у млекопитающих). Результаты показали, что в обоих случаях скорость эволюции для наружных участков молекулы почти в 10 раз выше, чем для области гемового «кармана» (табл. 7.1). Кроме того, известно [260], что два гистидиновых остатка, связанных в молекуле глобина с гемом, почти полностью инвариантны. Другими словами, за всю историю эволюции позвоночных, продолжавшейся около 500 млн. лет и включавшую очень много самых разных направлений, мутационные замены в этих двух положениях, по-видимому, почти не происходили. Редким исключением из этого правила является, вероятно, α -цепь глобина опоссума, аминокислотная последовательность которого была определена Стензелем [531] (см. гл. 4). Наглядной иллюстрацией связи между структурой и функцией этих молекул служит построенная Перутцем модель гемоглобинов [467]. Согласно более общей модели «масляной капли» для глобулярных белков [81], внутри молекулы находятся неполярные (гидрофобные) ами-

Таблица 7.1. Сравнение скорости эволюции для поверхности α - и β -цепей гемоглобинов и гемового «кармана». Скорости аминокислотных замен выражены в единицах $\times 10^{-9}$ /аминокислота за 1 год (в «полингах») [259]

Участок цепи	α -цепь	β -цепь
Поверхность	1,35	2,73
Гемовый «карман»	0,165	0,236

нокислоты, а на поверхности — полярные (гидрофильные). Активный центр локализован в «щели», и частота замещений аминокислот в этой области очень низка из-за жестких функциональных ограничений. С другой стороны, участки, расположенные на поверхности, обычно не очень существенны для функционирования или для поддержания четвертичной структуры, и скорость эволюции для них может быть гораздо выше в связи с тем, что доля селективно нейтральных мутаций здесь может быть значительно больше.

Можно было бы возразить, что различие в скорости эволюции для разных частей молекулы гемоглобина обусловлено различием в темпах мутирования, а не различием в селективных ограничениях. Обширные исследования различных вариантов гемоглобина в популяциях человека показали, однако, что это не так. Используя данные по описанным для популяций человека вариантам гемоглобина, Фогель [560] проанализировал распределение аминокислотных замещений в α - и β -цепях и обнаружил, что в α - и β -цистромах нет кодонов с повышенной тенденцией к мутированию. По-видимому, мутации, затрагивающие функционально «жесткий» гемовый «карман», происходят с такой же частотой, что и в функционально менее «жестких» участках на поверхности молекулы. В этом контексте особенно ценны данные о том, что в самых разных частях земного шара были обнаружены гемоглобины типа М, в которых один или два гистидина, примыкающие к гему, замещены тирозином, что вызывает специфическое заболевание. Эти варианты встречаются относительно редко, однако нельзя сказать, что они исключительно редки по сравнению с другими вариантами гемоглобина. Даже в пределах Японии были найдены все четыре типа вариантов гемоглобина М (подробнее см. [613]). Их назвали М Осака ($\alpha 58$, His $\rightarrow$ Tyr), М Ивате ($\alpha 87$, His $\rightarrow$ Tyr), М Куруме ($\beta 63$, His $\rightarrow$ Tyr) и М Акита ($\beta 92$, His $\rightarrow$ Tyr).

Причина инвариантности двух указанных гистидиновых остатков заключается в том, что мутации по соответствующим сайтам крайне неблагоприятны или близки к летальным в гомозиготном состоянии, хотя гетерозиготы при этом выживают. Известно, что около 30% молекул гемоглобина в крови гетерозигот по варианту М Ивате и другим гемоглобинам М аномальны (находятся в виде метгемоглобина), а остальные нормальны [509]. Таким образом, гетерозиготы, очевидно, селективно невыгодны, и эти мутации практически не могли распространиться среди позвоночных в ходе их эволюции. Аналогичным образом, согласно Хаяси и Стаматиянопулосу [175], предпоследний тирозиновый остаток в положении HC2 в гемоглобинах позвоночных также инвариантен. Исследование химических свойств вариантов Hb Бетесда и Hb Ренье, в которых этот тирозин замещен соответственно гистидином и цистеином (в β -цепи), пока-

зало, что в соответствии с моделью Перутца эти замены сопровождаются значительным нарушением процесса оксигенации. Таким образом, мутации по этим аминокислотным сайтам, несмотря на то, что они многократно возникали в ходе эволюции позвоночных, не смогли распространиться.

Другой интересный пример, иллюстрирующий тот факт, что менее важные в функциональном отношении участки молекул эволюционируют быстрее, — эволюция среднего сегмента (С) проинсулиновой молекулы. При образовании активного инсулина (состоящего из пептидов А и В) эта часть молекулы отщепляется. Известно, что она эволюционирует со скоростью примерно $2,4 \cdot 10^{-9}$ на аминокислотный сайт в год, т. е. в несколько раз быстрее, чем сам инсулин. На рис. 7.3 приведены значения скорости эволюции, рассчитанные по данным Джукса [199]; среднее различие между млекопитающими (за исключением морской свинки), указанное в табл. 2 этой работы, составляет 6% для пептидов А и В и 32% для пептида С. Нет сомнения, что в будущем удастся обнаружить и другие аналогичные примеры, так как образование молекул-предшественников, из которых получаются потом функционально активные формы, по-видимому, распространено довольно широко.

Мне хотелось бы привести еще один пример быстрой эволюции функционально менее значимой части белковой молекулы, рассмотренный в работе Барнарда и др. [18]. Эти авторы сообщили, что участок молекулы панкреатической рибонуклеазы, включающий аминокислоты с 15 по 24, эволюционирует с очень большой скоростью даже по сравнению с быстро эволюционирующим сегментом молекулы фибринопептида, и такую «гипервариабельность» можно объяснить тем, что этот участок совершенно неважен для ферментативной активности рибонуклеазы и поддержания структуры, необходимой для ее функционирования. В табл. 3 работы [18] сравнивается аминокислотный состав гипервариабельного сегмента фибринопептида А, участка 15—24 рибонуклеазы и пептида С проинсулина с аминокислотным составом «случайного белка». Не получив между ними никакого соответствия, авторы сделали вывод, что на все участки белковых молекул действует отбор и что истинно нейтральные мутации, как правило, в ходе эволюции не фиксируются. Я бы хотел обратить внимание на то, что аминокислотный состав данного белка для ограниченной группы видов может отличаться от состава «идеального случайного» белка просто в силу исторической случайности. Кроме того, селективное ограничение (отрицательный отбор) может в какой-то степени иметь место даже в гипервариабельных областях, но, как я уже подчеркивал, само существование селективного ограничения не противоречит теории нейтральности.

Чтобы выявить адаптивные замены на молекулярном уров-

не, Стеббинс [529] проанализировал данные по фибринопептидам. Воспользовавшись результатами работы [371] по аминокислотным последовательностям фибринопептидов А и В у парнокопытных, он оценил для каждого аминокислотного сайта частоты замен и коэффициенты корреляции Снита [520] между замещенной и новой аминокислотами. Стеббинс разделил замены на два типа: консервативные и радикальные. Если коэф-

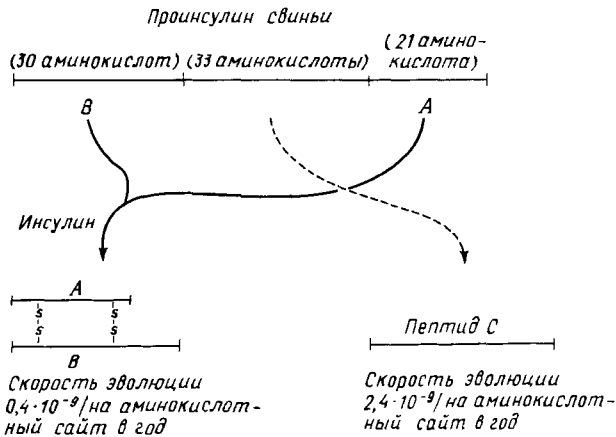


Рис. 7.3. Сравнение скоростей эволюции инсулина (пептиды А и В) и среднего сегмента молекулы проинсулина (пептид С).

фициент корреляции Снита для пары аминокислотных остатков превышал 0,6, такая замена считалась консервативной, в противном случае она рассматривалась как радикальная. Как считает сам Стеббинс, результаты этой работы совершенно четко указывали, что замены распределены по молекуле неравномерно, причем радикальные замены концентрировались в наиболее вариабельных сайтах, тогда как консервативные были распределены по тем сайтам, где замены происходили редко. Приняв во внимание сообщение о том, что аминокислотный остаток в положении 13 фибринопептида А влияет на скорость, с которой тромбин расщепляет эту молекулу, Стеббинс и Левонтин [529] постулировали, что радикальные замены в высоковариабельных участках (аминокислоты 12—19 в фибринопептиде А и 18—21 в фибринопептиде В) произошли в ходе эволюции парнокопытных в связи с каким-то адаптивным изменением. Эти авторы считают, что такие радикальные замены связаны с изменениями вспомогательных адаптивных свойств молекулы фибриногена и его производных и способствовали

приобретению самыми разными парнокопытными разнообразных физиологических и экологических особенностей.

Хотя такая трактовка кажется весьма правдоподобной, анализ фибринопептидов, сделанный Стеббинсом, ставит под сомнение некоторые ее положения. Рассмотрим вначале фибринопептид А. Известно, что фермент тромбин отщепляет фибринопептиды от фибриногена. Точка, в которой он отщепляет фибринопептид А, — это первый аргининовый остаток от NH_2 -конца цепи. Сравнение последовательностей фибринопептидов А показывает, что некоторые аминокислотные остатки почти или полностью инвариантны; таким остатком является, например, глицин в положении 4 (т. е. 4-й остаток перед аргинином, который тоже инвариантен). Валин в положении 2 высококонсервативен: анализ данных по парнокопытным выявил только одну замену ($\text{Val} \rightarrow \text{Ala}$). Коэффициент корреляции Снита (ϕ) для пары $\text{Val} \rightarrow \text{Ala}$ равен 0,675, т. е. эта замена консервативна. Глицин в положении 5 тоже высококонсервативен: те же данные выявляют здесь лишь одну замену ($\text{Gly} \rightarrow \text{Thr}$), причем радикальную, с коэффициентом корреляции $\phi = 0,396$.

По сравнению с этим участком молекулы положения 12—14 в высшей степени переменны. Например, в положении 13 (согласно данным по парнокопытным) встречаются 4 замены: $\text{Ala} \rightarrow \text{Val}$ ($\phi = 0,675$), $\text{Pro} \rightarrow \text{Ala}$ ($\phi = 0,553$), $\text{Val} \rightarrow \text{Pro}$ ($\phi = 0,473$) и $\text{Asp} \rightarrow \text{Ala}$ ($\phi = 0,156$). Согласно критерию Стеббинса, консервативна только первая замена ($\phi > 0,6$), остальные замены радикальны. Следовательно, замены в этом положении можно назвать преимущественно (на 75%) «радикальными». Однако и большинство других замен по этому критерию тоже становятся радикальными. Фактически из 190 коэффициентов корреляции для пар аминокислот, перечисленных в табл. 2 у Снита [520], только 16 больше 0,6, так что 174/190, или 92%, замен радикальны. Среднее для этих 190 коэффициентов корреляции составляет $\phi = 0,376$. Поэтому, вероятно, правильное в качестве граничной величины между консервативными и радикальными заменами выбрать коэффициент 0,4, а не 0,6. Таким образом, если мы будем считать замены с коэффициентами корреляции больше 0,4 консервативными, то замены в положении 13 будут в основном (на 75%) «консервативными» и лишь одна замена — радикальной ($\text{Asp} \rightarrow \text{Ala}$, $\phi = 0,156$). Кроме того, сравнение последовательностей разных А-пептидов четко показывает, что в этой переменной области довольно часто происходили вставки и делеции и восстановить исходную последовательность очень трудно: слишком велико число делеций и ни для одной из них не известна точная локализация. С фибринопептидом В ситуация еще хуже. Как правило, у фибринопептидов А животных переменным является участок с 14-го остатка по 19-й, а у фибринопепти-

дов В — с 9-го по 13-й [75]. Таким образом, результаты анализа, проведенного Стеббинсом, относительно природы аминокислотных замен с использованием фибринопептидов, неоднозначны и противоречат его утверждениям. Я полагаю, что эти и другие статистические соображения делают анализ Стеббинса сомнительным. Далее, хотя тот факт, что частоты аминокислотных замен распределены по пептидной цепи неслучайным образом, и свидетельствует о действии отрицательного отбора (совместимого с нейтралистской теорией), этого все же недостаточно, чтобы считать положительный дарвиновский отбор главным фактором, обуславливающим мутационные замены.

Как мы уже могли убедиться на нескольких примерах, противоречие между нейтралистами и селекционистами наиболее отчетливо проявляется в том, как они трактуют проблему быстрой эволюции молекул (например, фибринопептидов) или участков одной молекулы (например, пептида С проинсулина). Согласно теории нейтральности, последние функционально несущественны, поэтому очень многие мутации оказываются нейтральными и быстро накапливаются в результате случайного дрейфа. По мнению же сторонников селекционистской теории, быстро эволюционирующие молекулы или части молекул выполняют какие-то неизвестные функции и подвергаются быстрому адаптивному улучшению путем накопления множества благоприятных мутаций под действием положительного дарвиновского отбора. Какое из этих двух объяснений правильно, покажет будущее.

Здесь уместно сделать одно замечание. При обсуждении типов эволюции не следует путать скорость эволюции на уровне мутационных замен со скоростью изменения частоты какого-то одного мутантного аллеля. Первая (обозначаемая через k) измеряется путем подсчета числа аминокислотных (или нуклеотидных) замен, происшедших в течение длительного времени, скажем за десятки миллионов лет; вторая же (обозначаемая через Δp) определяется для коротких периодов времени, например для сотен и менее лет (как в случае с индустриальным меланизмом). К сожалению, здесь нередко наблюдается путаница, причем этим грешат даже некоторые генетики-популяционисты, когда они критикуют теорию нейтральности. Заметим, что для нейтрального аллеля степень изменения его частоты за поколение (Δp) может быть крайне малой, особенно если популяция достаточно велика, т. е. средняя величина изменения равна нулю, а стандартное отклонение — порядка $1/\sqrt{N_e}$. Скорость эволюции, выраженная через число мутационных замен k , не изменяется с изменением размера популяции N_e [см. уравнение 3.20]. С другой стороны, для благоприятного аллеля k значительно возрастает с увеличением размера популяции [см. уравнение (3.22)], а Δp при этом практически не изменяется.

Чтобы глубже понять механизм молекулярной эволюции, очень важно проанализировать зависимость между типом мутационных замен, с одной стороны, и четвертичной структурой и функцией молекулы — с другой. Мы все еще не можем описать в конкретных физико-химических терминах, как изменяются ограничения, налагаемые на молекулу, с появлением в ходе эволюции все новых и новых мутаций. Это относится в первую очередь к белкам, хотя очень полезной здесь может оказаться концепция Фитча о совместно варьирующих кодонах, или «коварионах». Согласно результатам статистического анализа, сделанного Фитчем и Марковицем [117], только 10% кодонов цитохрома *c* могут изменяться в ходе эволюции, но состав коварионов внутри молекулы изменяется все время. Так, оказалось [112], что доля коварионов составляет около 30% для α -цепи гемоглобина и почти 100% для фибринопептида А.

Очень интересен обнаруженный Фитчем факт, что если при расчете скорости аминокислотных замен брать за основу коварионы, то окажется, что цитохром *c*, α -цепь гемоглобина и фибринопептид А эволюционировали почти с одинаковой скоростью. С точки зрения теории нейтральности существование коварионов можно естественным образом объяснить, допустив, что они соответствуют аминокислотным сайтам, в которых мутации закрепляются путем случайного генетического дрейфа без изменения структуры и функций молекулы. Фитчу [113] удалось также оценить устойчивость ковариона, т. е. вероятность того, что кодон, принадлежащий к какой-то коварионной группе, останется таким же и после фиксации мутации в каком-то другом кодоне. Устойчивость по Фитчу оказалась менее 0,25, что предполагает довольно быструю реорганизацию коварионов внутри молекулы.

Вероятно, характер взаимодействия между аминокислотами в ходе эволюции определяется тем, насколько близко расположены они друг от друга в упакованной молекуле белка. В этой связи наиболее перспективным представляется исследование мутаций в триптофансинтетазе А у *E. coli*, проведенное Яновским и др. ([615]; см. также [569] и [614]). Этот белок состоит из 267 аминокислот. В положении 210 молекулы белка дикого типа находится глицин (см. рис. 7.4); исследование мутаций показало, что при замене глицина на серин активность фермента полностью сохраняется. С другой стороны, мутантный белок А46, у которого в положении 210 находится глутаминовая кислота, неактивен. Всестороннее исследование реверсий показало, что если происходит еще одна замена — Туг → Сус в положении 174 (на расстоянии 36 аминокислот от упомянутого выше глицина), — то активность фермента не изменяется (рис. 7.4). Однако единственная мутация Туг → Сус в положе-

нии 174 (при том, что в положении 210 находится, как в белке дикого типа, Gly) приводит к инактивации белка.

Аналогичным образом замена Gly→Val (мутант A187) в положении 212 сопровождается потерей активности, а в результате второй замены — Leu→Arg в положении 176 (опять отстоящем на 36 аминокислот от первичного мутационного сайта) — активность восстанавливается. Эти наблюдения позволили Яновскому и др. [615] предложить модель пространственного

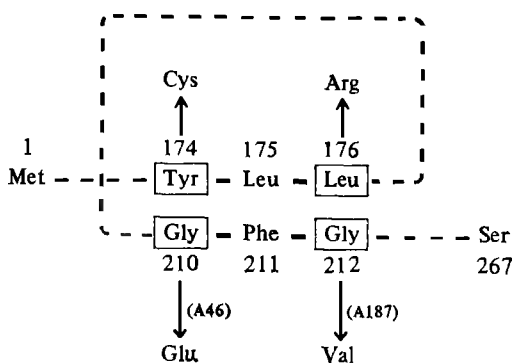


Рис. 7.4. Схема возможных взаимодействий между разными участками молекулы, полученная при исследовании обратных мутаций в триптофансинтетазе А [615].

расположения этих сайтов в упакованной молекуле белка (см. рис. 7.4).

Характер изменения селективных ограничений при замене одной аминокислоты на другую на самом деле должен быть более сложным, чем это представлено выше. Можно лишь с уверенностью сказать, что замена какой-то аминокислоты в одном положении часто влияет на целый набор сайтов в других частях молекулы, где возможны селективно нейтральные изменения.

Теперь я попытаюсь показать, что изменение селективных ограничений можно объяснить в более конкретных молекулярных терминах на примере транспортной РНК (тРНК), в частности рассматривая замены оснований в ее спиральных участках. Хорошо известно, что вторичную структуру всех тРНК можно представить в виде «клеверного листа», хотя при образовании четвертичной структуры молекула складывается более сложным образом [220, 274]. На рис. 7.5 изображена молекула дрожжевой тРНК^{Phe}, трехмерная структура которой [220, 221] напоминает букву L. Вообще говоря, молекула тРНК содержит четыре спиральных участка, обозначенных на рис. 7.5

через H_I , H_{II} , H_{III} и H_{IV} . Это акцепторный (связывающий аминокислоты), диридоуридиловый, антикодоновый и ТΨС стебли соответственно. В этих участках большинство оснований образуют уотсон-криковские пары, однако имеются одна или две неправильные пары, в частности в первом спиральном участ-

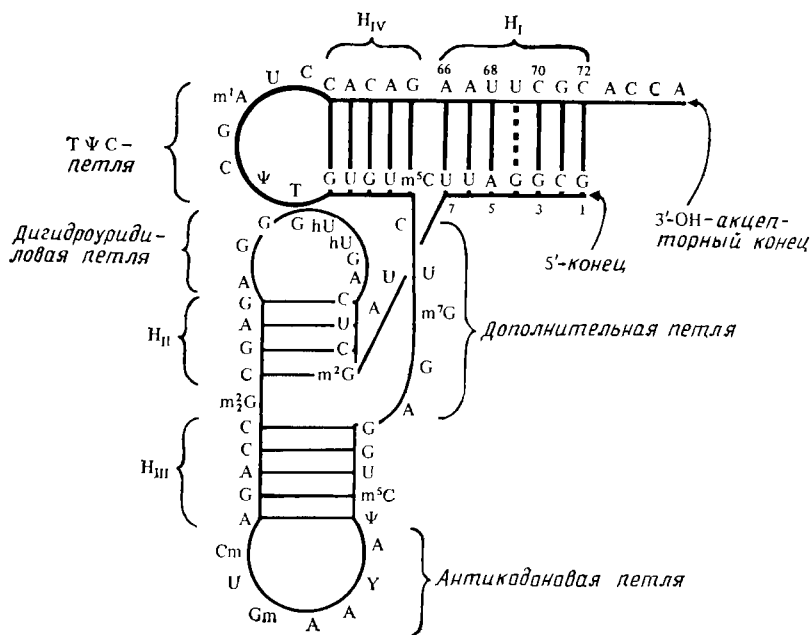


Рис. 7.5. Схематическое представление молекулы дрожжевой тРНК^{Phe} (по данным работы [221]).

ке. Было высказано предположение [199, 183], что единственное ограничение, налагаемое на структуру этих участков, заключается в том, что они должны находиться в форме двойной спирали. Это значит, что соответствующие основания могут претерпевать такие изменения в ходе эволюции, которые не сопровождаются нарушением спаривания по Уотсону — Крику; исключения составляют лишь некоторые сайты в этих участках. Согласно Холмквисту и др. [183], среди «неправильных» пар в спиральных участках пары GU (или UG), которые не влияют на спирализацию, встречаются гораздо чаще, чем другие. Фактически на их долю приходится свыше 60% от всех не уотсон-криковских пар. Интересно отметить также, что в спиральных участках пары GC, имеющие три водородные связи, встречаются гораздо чаще (69%), чем пары AU с двумя водородными связями. Это говорит о важности поддержания же-

стойкой двухцепочечной структуры. С другой стороны, в неспирализованных областях (петлях) присутствует гораздо меньше GC-пар (39%), что свидетельствует о необходимости сохранения открытой структуры, важной для узнавания соответствующих молекул. В петлях сосредоточено также большинство «инвариантных сайтов».

<i>тРНК^{Phe} E. coli</i>	<i>Дрожжевая тРНК^{Phe}</i>	<i>тРНК^{Phe} пшеницы</i>	<i>тРНК^{Lys} Drosophila</i>	<i>Дрожжевая тРНК^{Ala}</i>
1 G — C 72	G — C	G — C	G — C	G — C
C — G	C — G	C — G	C — G	G — C
3 C — G 70	C — G	G — C	C — G	G ... U
C — G	G ... U	G — C	C — G	C — G
5 G — C 68	A — U	G A	G ... U	G — C
G — C	U — A	G — C	G ... U	U U
7 A — U 66	U — A	A — U	C — G	G — C

Рис. 7.6. Наличие 0, 1 или 2 «неправильных» пар оснований в акцепторном стебле молекул тРНК.

Рассмотрим более детально спиральные участки, особенно связывающий аминокислоты акцепторный стебель (H_1), который состоит из семи пар оснований. На рис. 7.6 представлено несколько примеров, когда в этом стебле имеются 0, 1 и 2 неправильные пары. Кроме того, в табл. 7.2 перечислены частоты встречаемости таких пар. Среднее число неправильных пар на акцепторный стебель составляет примерно 0,59; среди них преобладают пары GU (0,49), а на долю всех остальных неправильных пар (GA, AC, UU) приходится 0,1. Более того, встречаемость неправильных пар других типов (не GU), по-видимому, не зависит от того, есть ли в этом стебле пара GU. Наблюдаемые частоты встречаемости разных типов неправильных пар хорошо согласуются с частотами, рассчитанными в предположении, что встречаемость таких пар среди семи парных сайтов статистически независима.

Мы можем представить себе процесс эволюционных изменений в этой части молекулы как последовательность сменяющих друг друга процессов образования и разрушения уотсон-криковских пар. Как указывала Ота [435], первое разрушение, вероятно, было лишь слегка неблагоприятным, и мутация, вызвавшая такое изменение, могла распространиться в популяции путем случайного дрейфа. Вероятно, образование одной неправильной пары, особенно если это пара GU, сопровождается настолько малыми нарушениями, что является практически нейтральным с точки зрения естественного отбора. Как считал Джукс [199], пара GU представляет собой наиболее естественную переходную ступень от GC к AU или наоборот. С другой стороны, неправильные «не GU»-пары несомненно

Таблица 7.2. Частоты встречаемости разных «неправильных» пар в акцепторном стебле тРНК. Наблюдаемые частоты были рассчитаны по данным о последовательности 177 тРНК, представленным в обзоре [126], за исключением чувствительной к температуре тРНК *E. coli*. В этой таблице «не GU» означает все другие неправильные пары (например, CA и UU). Ожидаемые частоты рассчитаны в предположении, что частоты встречаемости неправильных пар среди семи парных сайтов статистически независимы

Неправильные пары	Частота, %	
	наблюдаемая	ожидаемая
Отсутствуют	55,1	54,4
1GU	25,0	28,6
2GU	9,7	7,1
3GU	0,6	1,0
1 «не GU»	6,8	5,8
1GU+1 «не GU»	2,8	3,1
Всего	100,0	100,0

ослабляют структуру, и хотя одна такая пара иногда может оказаться допустимой, мутации, приводящие к более чем одной подобной «неточности», элиминируются из-за слишком большого неблагоприятного эффекта; молекула может «выдержать» появление новой мутации только в том случае, если уже имеющаяся неправильная пара «маскируется» какой-либо мутантной заменой. Таким образом, несмотря на наличие в акцепторном стебле семи пар оснований (четырнадцать нуклеотидных сайтов), реальное число варьируемых сайтов в любой момент на самом деле меньше.

Одзечи и др. [463] использовали нонсенс-супрессорные мутанты *E. coli* для того, чтобы показать, что правильное функционирование этого участка тРНК зависит от наличия уотсон-криковских пар (а иногда пары GU), а не от природы самих оснований. Например, в мутанте su^{+2} (тРНК₂^{Gln}), у которого «amber»-кодон (терминирующий кодон UAG) транслируется в глутамин, пары оснований расположены в акцепторном стебле в следующем порядке:

7	6	5	4	3	2	1
A	U	G	G	G	G	U
.	.	.	.	.	.	.
U	A	C	C	C	C	A

Было обнаружено, что если А в положении 7 заменяется на С (такой мутант обозначается С7), молекула становится «неполноценной». Мутант G65, в котором основание U (спаренное с А в положении 7) заменено на G, также неполноценен.

Однако двойной мутант C7-G65, у которого пара AU заменена на GC, функционирует нормально. Одзэки и др. [463] рассматривают много интересных примеров такого рода.

7.3. Синонимичные и другие «молчащие» замены

В гл. 4 мы уже говорили о высокой частоте синонимичных замен нуклеотидов в ходе эволюции. Мы отмечали также, что поскольку синонимичные замены подвергаются действию естественного отбора в гораздо меньшей степени, чем несинонимичные, то наблюдаемые для синонимичных замен высокие частоты подтверждают теорию нейтральности.

положения аминокис- лот	2	④	⑤	10	11	19	②1	33	36	42	47	⑤0		
<i>β</i> -цепь гло- бина человека	His CAC	Thr ACU	Pro CCU	Ala GCC	Val GUU	Asn AAC	Asp GAU	Val GUG	Pro CCU	Phe UUU	Asp GAU	Thr ACU		
<i>β</i> -цепь глоби- на кролика	His CAU	Ser UCC	Ser AGU	Ala GCG	Val GUC	Asn AAU	Glu GAA	Val GUU	Pro CCA	Phe UUC	Asp GAC	Ser UCU		
	⑤1	⑤2	⑤6	57	66	68	⑥9	71	⑦3	74	⑦6	82	86	⑧7
	Pro CCU	Asp GAU	Gly GGC	Asn AAC	Lys AAA	Leu CUC	Gly GUU	Phe UUU	Asp GAU	Gly GGC	Ala GCU	Lys AAG	Ala GCC	Thr ACA
	Ala GCA	Asn AAU	Asn AAC	Asn AAU	Lys AAG	Leu CUG	Ala GCU	Phe UUC	Glu GAG	Gly GGU	Ser AGU	Lys AAA	Ala GCU	Lys AAG
	90	111	①12	①15	117	123	124	①25	132	138	142	144	145	
	Glu GAG	Val GUU	Cys UGU	Ala GCC	His CAC	Thr ACC	Pro CCA	Pro CCA	Lys AAA	Ala GCU	Ala GCC	Lys AAG	Tys UAU	
	Glu GAA	Val GUU	Ile AUU	Ser UCU	His CAU	Thr ACU	Pro CCU	Gln CAG	Lys AAG	Ala GCC	Ala GCU	Lys AAA	Tyr UAC	

Рис. 7.7. Схема, иллюстрирующая преобладание синонимичных замен, которое было выявлено при сравнении генов β -цепей глобинов человека и кролика. Положения, в которых последовательности β -цепей содержат разные аминокислоты, отмечены числами, заключенными в кружок. Различающиеся основания в гомологичных сайтах заключены в рамку. Указаны только те кодоны, которые различаются по крайней мере одним основанием.

На рис. 7.7 представлены данные, иллюстрирующие преобладание синонимичных замен, которое было выявлено при сравнении последовательностей участка генов, детерминирующих синтез β -цепи глобинов человека [297] и кролика [558]. Изображены только те пары кодонов, которые различаются хотя бы одним основанием. Различающиеся основания в гомологичных сайтах заключены в рамку. Число синонимичных замен для оснований превосходит число различий по аминокислотам

Таблица 7.3. Возможные различия между гомологичными нуклеотидными последовательностями

	Транзния	Трансверсия	
Последовательность 1	U C A G	U A C G	U G C A
Последовательность 2	C U G A	A U G C	G U A C
Частота	P	Q	R

ветствующую величину для пар UG, GU, CA и AC. Таким образом, $Q+R$ представляет собой частоту гомологичных сайтов, различающихся по типу трансверсии.

Тогда эволюционное расстояние между двумя последовательностями, выраженное через число замен оснований, можно оценить так:

$$K = -(1/4) \ln [(1-2P-2Q)(1-2P-2R)(1-2Q-2R)]. \quad (7.4)$$

Это сумма трех компонент:

$$2\alpha T = -(1/4) \ln [(1-2P-2Q)(1-2P-2R)/(1-2Q-2R)], \quad (7.4a)$$

$$2\beta T = -(1/4) \ln [(1-2P-2Q)(1-2Q-2R)/(1-2P-2R)], \quad (7.4b)$$

$$2\gamma T = -(1/4) \ln [(1-2P-2R)(1-2Q-2R)/(1-2P-2Q)]. \quad (7.4b)$$

В частном случае, когда два типа замен по типу трансверсии встречаются с одинаковой частотой, т. е. $\gamma = \beta$, эта формула сводится к

$$K = -(1/2) \ln [(1-2P-Q') \sqrt{(1-2Q')}], \quad (7.5)$$

где $Q' = Q+R$ — общая доля замен по типу трансверсии. Это согласуется с уравнением (4.14). Если время дивергенции T известно, то частоту замен оснований на сайт за год можно оценить как

$$k = K/2T. \quad (7.6)$$

Известно, что значительная часть замен оснований по третьему положению кодонов является синонимичной, поэтому интересно оценить «синонимичную» компоненту числа замен оснований по этому положению. Она выражается формулой

$$K_s' = -\left(\frac{1}{4}\right) \ln [(1-2P-2Q)(1-2P-2R)]. \quad (7.7)$$

В частном случае, когда $\gamma = \beta$, эта формула сводится к (4.16). Кроме того, можно получить формулу для дисперсии этих оценок. Более подробно см. об этом в работе [244].

В табл. 7.4 представлены результаты, полученные с помощью этих формул для различных случаев. В первой строке

Таблица 7.4. Эволюционные расстояния (со стандартными ошибками) на нуклеотидный сайт, рассчитанные методом 3ST. K_i ($i=1, 2, 3$) — число замен оснований в положении i кодона, по которым различаются две сравниваемые последовательности; K_S' — синонимичная компонента числа замен по третьему положению кодона

	Эволюционное расстояние на нуклеотидный сайт			
	K_1	K_2	K_3	K'_S
Предшественник гормона роста: человек и крыса	$0,26 \pm 0,04$	$0,18 \pm 0,03$	$0,53 \pm 0,07$	$0,44 \pm 0,07$
Предшественник проинсулина I: человек и крыса				
цепи A+B (инсулин)	$0,04 \pm 0,03$	$0,00^1)$	$0,46 \pm 0,12$	$0,38 \pm 0,12$
пептид С	$0,18 \pm 0,06$	$0,27 \pm 0,10$	$0,95 \pm 0,46$	$0,77 \pm 0,51$
β -Цепь глобинов: человек и кролик	$0,06 \pm 0,02$	$0,06 \pm 0,02$	$0,28 \pm 0,05$	$0,25 \pm 0,05$
β -Цепь глобинов: кролик и мышшь	$0,16 \pm 0,03$	$0,13 \pm 0,03$	$0,43 \pm 0,07$	$0,36 \pm 0,07$
α -Цепь глобинов: кролик и мышшь ²⁾	$0,12 \pm 0,03$	$0,12 \pm 0,03$	$0,54 \pm 0,09$	$0,47 \pm 0,09$
<i>S. purpuratus</i> и <i>P. miliaris</i> :				
гистон H2В	$0,09 \pm 0,03$	$0,02 \pm 0,01$	$0,48 \pm 0,10$	$0,43 \pm 0,10$
гистон H3	$0,008 \pm 0,008$	$0,008 \pm 0,008$	$0,47 \pm 0,08$	$0,41 \pm 0,08$

¹⁾ Среди 51 кодона изменений не наблюдалось.

²⁾ Ген α -I мыши (из [421]).

таблицы сравниваются нуклеотидные последовательности генов, детерминирующих синтез предшественника гормона роста (пресоматотропина) человека [330] и крысы [505]. Значения K для первого, второго и третьего положений кодонов равны 0,26, 0,18 и 0,53 соответственно, а для синонимичного компонента по третьему положению — 0,44. Поскольку человек и крыса, вероятно, дивергировали в мезозое, т. е. около 80 млн. лет назад, мы можем принять $T=8 \cdot 10^7$. Тогда скорость эволюции за год по трем положениям кодона составит $k_1=1,6 \cdot 10^{-9}$, $k_2=1,1 \cdot 10^{-9}$ и $k_3=3,3 \cdot 10^{-9}$, а $K_S'=2,8 \cdot 10^{-9}$. Что касается несинонимичных замен, то у предшественника гормона роста они, по-видимому, происходили немного чаще, чем у гемоглобинов.

Во второй и третьей строках табл. 7.4 сравниваются гены предшественника проинсулина человека [19, 535] и крысы [55]; интересно, что скорости эволюции для цепей A+B инсулина и пептида С по первому из двух положений кодона сильно различаются: судя по скорости аминокислотных замен, инсулин эволюционировал примерно в 10 раз медленнее, чем пептид С проинсулина. Если сравнивать синонимичные компоненты (K_S'),

то скорость эволюции инсулина окажется лишь немного ниже, чем для пептида С. Более того, значения K_s' для предшественника проинсулина и предшественника гормона роста примерно одинаковы (в случае дивергенции человек — крыса). Значения K , приведенные в четвертой, пятой и шестой строках таблицы, были рассчитаны с использованием данных по генам β -цепи глобинов человека [328], кролика [92], мыши [283], α -цепи глобина кролика [177] и α -1-цепи глобина мыши [421].

Если мы сравним эти значения с соответствующими данными, представленными в табл. 4.8 (см. строки 3, 4 и 6), то обнаружим, что в отношении α - и β -цепей гемоглобинов для этих пар животных рассматриваемая модель дает почти такие же оценки, как и в случае более простой модели (рис. 4.9), предложенной в гл. 4. Время дивергенции кролик — мышь, как и человек — кролик, можно считать равным 80 млн. лет ($T=8 \cdot 10^7$).

В данном контексте представляет интерес вопрос о дивергенции интронов. Сравним малые интроны (интрон 1) человека и кролика. Два этих гомологичных интрона будут достаточно хорошо соответствовать друг другу, если в интрон кролика «вставить» три «пропуска» (насчитывающих до четырех нуклеотидов). Исключая из рассмотрения эти пропуски, а также два первых и два последних нуклеотида (которые должны быть инвариантными согласно правилу GT/AG для сплайсинга; см. [421]), найдем $P=16/122$, $Q=8/122$, $R=11/122$. Подставив эти значения в формулу (7.4) и воспользовавшись формулой для стандартной ошибки (см. [244]), получим расстояние на сайт $K=0,36 \pm 0,053$. Его можно сравнить с соответствующей оценкой для синонимичной компоненты $K_s'=0,25$ (см. строку 4 в табл. 7.4). Учитывая, что $2/3$ замен в положении 3 синонимичны, мы можем оценить уточненное «синонимичное» расстояние: $K_s=K_s' (2/3)$, откуда $K_s=0,38$. Таким образом, скорость мутационных замен в малом интроне примерно равна скорости синонимичных замен.

В двух последних строках табл. 7.4 приведены эволюционные расстояния между двумя видами морских ежей, *S. purpuratus* и *P. miliaris*, оцененные исходя из данных по аминокислотным последовательностям гистонов H2B и H3 (использовались данные работ [536] и [503]). Эти два вида морских ежей принадлежат к разным семействам, дивергировавшим от 60 до 160 млн. лет назад [213]. Полученные значения K_1 и K_2 показывают, что, если судить по частоте замен аминокислот, эти гистоны эволюционировали очень медленно; кроме того, скорость эволюции гистона H3 составляет лишь $1/7$ скорости для гистона H2B. Однако если рассматривать синонимичную компоненту скорости замен по третьему положению кодонов, то эти гистоны оказываются очень близкими. Кроме того, скорость синонимичных замен у них не очень отличается от соответству-

ющей скорости для предшественника гормона роста, инсулина и α - и β -цепей глобинов. Приняв время дивергенции этих видов морских ежей равным 10^8 лет, мы получим $k_s' = 2,2 \cdot 10^{-9}$ для гистона H2B и $k_s' = 2,1 \cdot 10^{-9}$ для гистона H3.

Еще более консервативным, чем гистон H3, и, видимо, самым консервативным из всех белков является гистон H4. Как отмечалось в разд. 7.2, скорость замен аминокислот у гистона H4 по оценкам равна $k_{aa} = 0,008 \cdot 10^{-9}$ на аминокислотный сайт

<i>L. pictus</i> мРНК	GAU	AAC	AUC	CAA	GGA	AUA	ACU	AAA	CCG	GCA
<i>S. purpuratus</i> мРНК	GAU	AAC	AUC	CAA	GGC	AUC	ACC	AAG	CCU	GCA
Гистон H4 Аминокислотная последовательность	Asp	Asn	Ile	Gln	Gly	Ile	Thf	Lys	Pro	Ala
	24	25	26	27	28	29	30	31	32	33

Рис. 7.9. Сравнение части нуклеотидных последовательностей мРНК гистона H4 у морских ежей *S. purpuratus* и *L. pictus*; приведена также соответствующая аминокислотная последовательность гистона H4 быка. Диаграмма построена по данным работ [148, 149].

в год. Примечательно, что даже в таком высококонсервативном белке в ходе эволюции происходит достаточно много синонимичных замен; это видно из рис. 7.9, где сравнивается часть нуклеотидных последовательностей мРНК для гистона H4 двух видов морских ежей — *Lytechinus pictus* и *Strongylocentrotus purpuratus*. Используя данные работы [149], я оценил эволюционное расстояние между этими видами по третьему положению кодонов для гистона H4 и получил $K_3 = 0,44 \pm 0,163$ [236]. Как считают авторы работы [149], эти морские ежи дивергировали примерно $6 \cdot 10^7$ лет назад. Следовательно, если принять $T = 6 \cdot 10^7$, мы получим $k_3 = (3,7 \pm 1,4) \cdot 10^{-9}$. Эта величина тоже сравнима со скоростью замен по третьему положению кодона для ряда других белков.

В качестве последнего примера высококонсервативного белка я хотел бы рассмотреть тубулин — белок микротрубочек, обладающий субъединичной структурой. Недавно была определена нуклеотидная последовательность, соответствующая С-концевому участку α -тубулина крысы [132], и появилась возможность сравнить ее с соответствующей последовательностью для α -тубулина курицы [554]. Я использовал данные рис. 2 из работы Гинзбурга и др. [132]. Исключив кодон терминации, можно сравнить 145 кодонов. Оказалось, что С-концевые участки указанных тубулинов различаются только по одной аминокислоте, т. е. число аминокислотных замен составляет $K_{aa} = 0,0069$. С другой стороны, синонимичные изменения затрагивают 51 кодон, причем в одном из них эти изменения происходят в первом и третьем положениях (AGA вместо CGC; оба кодируют Arg).

По третьему положению кодона частоты трех типов замен составляют $P=33/145$, $Q=6/145$, $R=12/145$. Отсюда получаем оценку числа замен оснований, по которым различаются α -тубулины крысы и курицы, $K_3=0,51\pm 0,08$ для третьего положения кодонов. Синонимичная компонента в данном случае составляет $K_s'=0,44\pm 0,08$. Интересно отметить, что эти величины не очень отличаются от соответствующих величин ($K_3=0,64\pm 0,11$ и $K_s'=0,53\pm 0,10$), полученных при сравнении β -цепей глобина курицы и кролика (см. табл. 4.8), несмотря

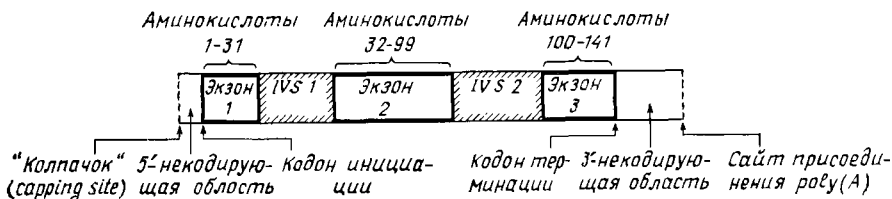


Рис. 7.10. Структура гена α -цепи глобина (α -2-гена) человека (по данным работы [475]). Структура гена β -цепи глобина человека аналогична, за исключением того, что второй интрон имеет большую протяженность (573 пары оснований в гене β -цепи глобина кролика; Ouyen et al., 1979). Символом IVS обозначены интроны. Длина интрона IVS1 — 117 пар оснований, интрона IVS2 — 140 пар оснований.

на огромное различие скоростей несинонимичных замен между β -цепью глобина и α -тубулином.

Тот факт, что скорости синонимичных замен не только выше, чем несинонимичных, но и примерно одинаковы для разных генов, был установлен также в результате всесторонних исследований Мияты и др. [365], которые сравнили около 50 пар гомологичных нуклеотидных последовательностей разных генов. Для средней скорости синонимичных замен они получили оценку $5,1 \cdot 10^{-9}$ на нуклеотидный сайт в год. Грубая оценка, полученная мною из данных табл. 7.4 для синонимичной компоненты скорости замен в третьем положении кодонов, составляет $k_s'=2,6 \cdot 10^{-9}$ на сайт в год. Если мы предположим, что четыре типа оснований встречаются с одинаковой частотой, то синонимичными по третьему положению кодона окажутся около $2/3$ нуклеотидных замен. Следовательно, скорость синонимичных замен в ходе эволюции можно оценить, разделив K_s' на $2/3$, что дает $k_s=3,9 \cdot 10^{-9}$ за год. Эта величина не очень отличается от оценки Мияты и др. — $5,1 \cdot 10^{-9}$.

Мията и др. [365] попытались определить также скорость эволюции некодирующих участков генов. Они обнаружили, что последовательности интрона и 5'-участка 3'-некодирующей области дивергировали почти в такой же степени, как и синонимичные кодоны (термины, использующиеся при описании ге-

нов эукариот, приведены на рис. 7.10). Напротив, другие сегменты, включающие 5'-некодирующую область и 3'-участок 3'-некодирующей области, строго консервативны: степень их дивергенции примерно вдвое меньше, чем для синонимичных сайтов. Это вполне понятно, поскольку здесь находятся различные сигнальные последовательности, контролирующие транскрипцию и модификацию мРНК, что налагает строгие ограничения на возможные мутационные замены оснований. В число таких сигнальных последовательностей входят ТАТА («tar-tar» box) для начала транскрипции, так называемый «колпачок» («capping»), ААТААА-сигнал для присоединения ро-ly(A) и ряд других (см., например, [298, 537]).

Все эти данные легко объяснить в рамках теории нейтральности, согласно которой большинство мутационных замен у данного вида обусловлено случайной фиксацией селективно нейтральных (т. е. селективно эквивалентных, но не обязательно идентичных с функциональной точки зрения) мутаций, а не положительным дарвиновским отбором. Хотя существование благоприятных мутаций не вызывает сомнений, в этой теории предполагается, что они происходят достаточно редко и ими можно пренебречь. Согласно теории нейтральности, вероятность того, что мутация будет селективно нейтральной (т. е. не вредной), тем больше, чем в меньшей степени она сказывается на структуре и функции молекулы. В той мере, в какой все мутации являются селективно нейтральными, скорость эволюции на сайт k равна общей скорости мутирования ν на сайт. Вероятно, синонимичные мутации близки к нейтральным [236], поэтому скорость синонимичных замен на сайт у разных молекул примерно одинакова. Конечно, это равенство частот весьма приблизительное, так как не исключено все же (см. разд. 7.5), что существует селективное ограничение, связанное с «неслучайным» использованием синонимичных кодонов. Однако степень такого селективного ограничения, по-видимому, достаточно мала. Аналогичным образом можно объяснить быстрое эволюционное изменение интронов (см. гл. 4).

Противоположный взгляд был высказан Перье и др. [466]. Эти авторы считают, что движущей силой фиксации является положительный естественный отбор, действующий на какую-то часть несинонимичных изменений («замещений»), и что такая селективная фиксация распространяется и на нейтральные изменения (в том числе на изменения по молчащим сайтам), которые накопились в этих участках ДНК. Другими словами, для объяснения фиксации синонимичных замен они привлекли эффект так называемого «попутного транспорта» (hitch-hiking).

В своей работе [244] я уже отмечал, что таким образом нельзя объяснить имеющиеся данные, не отвергая при этом основные принципы популяционной генетики. На самом деле с

помощью такого механизма не может происходить достаточно быстрая фиксация нейтральных мутаций, так как сами селективные изменения накапливаются очень медленно. Рассмотрим, например, ген гистона H4. Скорость замещений здесь крайне мала, скорость же синонимичных замен оснований сравнима с таковой для замещений в фибринопептидах — очень быстро эволюционирующих молекулах.

Эту проблему можно проанализировать еще детальнее. Так как эффект «попутного транспорта» в случае фиксации сцепленных аллелей распространяется лишь на небольшое расстояние от подвергающегося отбору гена, мы рассмотрим в качестве некоего генного локуса малый сегмент ДНК, в котором кроссинговер происходит так редко, что им можно пренебречь. Предположим, что в этом локусе появился один благоприятный мутантный аллель. Для того чтобы такая поддерживаемая отбором мутация могла способствовать фиксации других, неотбираемых (нейтральных) мутаций путем «попутного транспорта», копия гена, в которой появилась эта благоприятная мутация, должна одновременно содержать и ряд нейтральных мутаций. Кроме того, чтобы скорость замен нейтральных мутаций на сайт в m раз превышала скорость замен, обусловленных мутациями, на которые действует положительный отбор (в данном случае несинонимичных замен), каждая копия гена в популяции в среднем должна содержать m нейтральных мутаций безотносительно к тому, оказываются среди них благоприятные мутации или нет. Коэффициент m должен быть очень большим: для гистона H4 он, вероятно, равен 1000 или более. С другой стороны, если бы каждая копия гена содержала большое число нейтральных мутаций, то соответствующие (гомологичные) гены у разных особей настолько сильно отличались бы друг от друга, что такого понятия, как видоспецифичная нуклеотидная последовательность определенного гена, просто не существовало бы. Иными словами, гомологичные последовательности у всех особей данного вида были бы абсолютно разными. Все это совершенно не соответствует действительности.

Гипотеза «попутного транспорта» не может объяснить и еще один факт (уже обсуждавшийся в этом разделе), наблюдавшийся при исследовании генов разных белков, а именно: скорости синонимичных замен в ходе эволюции у разных белков не только велики, но и примерно одинаковы, тогда как скорости замен несинонимичных могут существенно различаться.

Джукс [204], проделавший полный анализ встречаемости синонимичных замен у разных организмов, отмечает, что половину нуклеотидных замен, происшедших в ходе дивергенции генов у животных, бактерий и вирусов, составляют молчащие замены. Он тоже считает, что эти изменения есть результат слу-

чайного дрейфа, оперирующего с нейтральными или почти нейтральными изменениями (см. также [205]).

Важнейшая функция ДНК состоит в самовоспроизведении, но у нее есть и другие свойства — способность к накоплению мутаций и случайных дупликаций.

Представление о «гене как основе жизни», которого долгое время придерживался Мёллер [394], было значительно расширено и распространилось на признаки, связанные с поведением; появилось целое новое направление — социобиология [583]. Один из аспектов представлений социобиологов нашел отражение в названии книги Доукинса «Эгоистичный ген» [73]. Перефразируя его, Оргел и Крик [461] предложили концепцию «эгоистичной» ДНК (см. также работу [88]). Согласно этим авторам, эгоистичная ДНК — это сегмент ДНК, который обладает очень малым фенотипическим эффектом или не обладает им вообще, но тем не менее широко представлен у данного вида благодаря интенсивной репликации внутри генома. Единственная задача такой ДНК — самосохранение, поэтому ее можно рассматривать как «безусловного паразита». К этому классу относятся нуклеотидные последовательности у эукариот с большим и средним числом повторов, а также подвижные элементы генома типа «соря» у *Drosophila*.

Концепция эгоистичной ДНК вызвала очень большой интерес и многократно обсуждалась (см., например, работу [462]). Однако в таких дискуссиях часто недооценивался один момент: для распространения «эгоистичного» элемента по популяции данного вида недостаточно одной лишь способности его к сверхрепликации внутри генома. Для этого должны произойти транспозиция в другой геном (хромосомный набор) в пределах диплоидной особи и распространение по другим особям путем полового процесса.

Это означает, что случайный генетический дрейф должен играть очень важную роль в судьбе эгоистичной ДНК в популяции, так как по определению такая ДНК не оказывает никакого специфического влияния на фенотип и, следовательно, в дарвиновском смысле селективно нейтральна. Поэтому весьма целесообразно было бы рассмотреть эволюцию эгоистичной ДНК с позиций популяционной генетики. Начало этому направлению было положено Отой и Кимурой [438, 449]. Эгоистичная ДНК была разделена на два класса: тандемно повторяющиеся последовательности и рассеянные по геному повторяющиеся семейства. Однако на самом деле эти классы не могут быть полностью независимыми, как следует из недавней работы Чайлдса и др. [47], сообщивших о новом классе генетических элементов, которые были названы орфонами. Орфоны — это рассеянные единичные элементы, которые, как считается, переместились из тандемно повторяющихся мультигенных семейств в новые ме-

ста (подобно сиротам, потерявшим свои семьи). Чайлдс и др. детально проанализировали орфоны гистонового гена морского ежа *L. pictus* и обнаружили, что в геноме каждого ежа имеется более 50 генных орфонов. Кроме того, они нашли, что локусы орфонов в популяциях полиморфны. По-видимому, в ходе эволюции орфоны отделились от родительских семейств относительно недавно и обладают свойствами, отличающими их от обычных транспозонов. Популяционно-генетические эффекты таких элементов пока неизвестны.

7.4. Быстрые эволюционные изменения псевдогенов

Псевдоген — это сегмент ДНК, имеющий четкую гомологию с каким-либо известным функционирующим геном, но утративший вследствие мутационных изменений способность к детерминации функционального продукта [474]. Иногда его называют «мертвым геном», однако против этого термина имеются возражения (см., например, [557]): высказывалось мнение, что псевдоген может выполнять хотя и незначительную, но важную функцию, например контролировать экспрессию обычных генов.

По-моему, гораздо разумнее считать псевдогены действительно мертвыми; возникшие путем генной дупликации и затем утратившие свои функции вследствие мутаций, они освободились от контроля отрицательного отбора. В дальнейшем они накапливали различные мутационные изменения с максимальной скоростью, которая только возможна при мутационном давлении и случайном дрейфе. Если все это на самом деле так, то степень их эволюционных изменений должна быть такой, как это предсказывает теория нейтральности. При сравнении псевдогенов с их нормальными партнерами обнаруживается немало замечательных свойств. В них выявляются не только происшедшие с большой скоростью замены оснований, но и весьма частые делеции и вставки. В отличие от «консервативного» характера изменений, который свойствен эволюции многих обычных генов, замены оснований по первому и второму положениям кодона в псевдогенах встречаются так же часто, как и замены по третьему положению. В конце гл. 4 мы уже бегло обсуждали эти особенности для глобиноподобного α -3-гена [421].

Нередко изменения, сопровождающиеся утратой функции, происходят и в различных сигнальных последовательностях в некодирующем участке; эти последовательности играют важную роль в контроле транскрипции и различных посттранскрипционных модификаций. Например, у мыши имеются два псевдо- α -глобиновых гена, обозначаемых α -3 и α -4 [298]. В одном из них (α -3) промотор ТАТА заменен на TAGA. Из экспериментов Василика и др. [568] следует, что у α -3-гена должна быть нарушена важнейшая функция инициации транскрипции. Эти

же авторы показали, что при замене в гене кональбумина третьего основания в последовательности ТАТА (Т на G) эффективность его специфической транскрипции *in vitro* резко снижается. По данным работы [532], делеция второго динуклеотида ТА в последовательности ТАТАААА (положения —30 ~ —24) в гене фиброина шелка у *Bombyx mori* приводит к значительному подавлению транскрипции фибринового гена. В другом псевдогене мыши (α -4) изменена сигнальная последовательность для poly(A) — от нормальной ААТАА на аномальную ААТАС.

Позже мы еще раз вернемся к псевдогену α -3, а теперь вкратце рассмотрим весьма интересную группу изменений, обнаруженных в кодирующей области псевдогена α -4. Согласно Ледеру и др. [298], ген α -4 содержит точковые мутации, в результате которых гистидин в 58-м положении аминокислотной последовательности заменяется на тирозин. Это соответствует гемоглобину М Бостон, вызывающему у человека метгемоглобинемию. Кроме того, в гене α -4 произошли изменения, которые сдвигают по фазе нормальный терминирующий кодон, так что при синтезе глобина считываются 40 дополнительных кодонов аналогично тому, как это происходит в случае гемоглобина Констант Спринг у человека. Так как у нормальных мышей таких аномальных гемоглобинов не обнаружено, ясно, что ген α -4 не экспрессируется. Эти изменения вместе с ранее упоминавшимися говорят о том, что псевдогены постоянно дезинтегрируются, превращаясь, по словам Оно [430], в «хлам» («junk»).

Хотя и раньше были выполнены отдельные расчеты [241, 475], показавшие, что псевдогены эволюционировали быстрее (если говорить о частоте замен оснований), чем их нормальные «партнеры», первыми, кто провел тщательный анализ скорости эволюции псевдогенов, были все же Мията и Ясунага [364]. Они использовали данные, опубликованные Ваниным и др. [557], по нуклеотидной последовательности мышинного псевдогена, родственного α -глобиновому гену и обозначенного $\psi\alpha$ 30.5, который в значительной степени эквивалентен мышинному α -3-гену Нисиоки и др. [421]. Этот ген обладает замечательной особенностью: у него полностью отсутствуют два интрона, имеющиеся во всех функциональных α - и β -глобиновых генах (т. е. по структуре он полностью соответствует α -глобиновой мРНК после процессинга). Мията и Ясунага [364] сравнили последовательности этого псевдогена и α -глобиновых генов мыши и кролика. Оказалось, что по каждому функциональному или структурному блоку, по которым можно провести такое сравнение, мышинный ген $\psi\alpha$ 30.5 (сокращенно М $\psi\alpha$) всегда гораздо ближе к нормальному α -глобиновому гену взрослой мыши (М α), чем к гену кролика (R α). Единственное исключение составляет 3'-часть 3'-некодирующей области, что можно объ-

яснить случайными и/или структурными изменениями, обусловленными неравным кроссинговером в этом участке $M\psi\alpha$ (структура α -глобинового гена представлена на рис. 7.10). Такой результат предполагает, что этот псевдоген образовался из нормального α -глобинового гена путем генной дупликации, происшедшей в геноме мыши после дивергенции линий мыши и кролика. (Согласно недавнему сообщению Людерса и др. [317], псевдогены мыши $\alpha\psi 3$ и $\alpha\psi 4$, обозначенные выше $\alpha-3$ и $\alpha-4$, локализованы в хромосомах 15 и 17 соответственно, тогда как

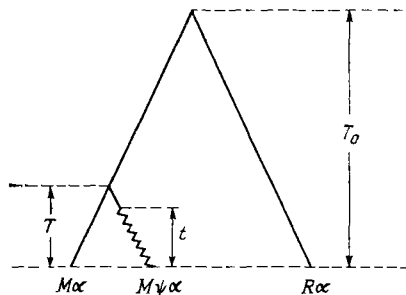


Рис. 7.11. Филогенетические связи, предложенные Миятой и Ясунагой [364] для анализа псевдо- α -глобинового гена мыши ($M\psi\alpha$). Буквами Ma и Ra обозначены нормальные α -глобиновые гены взрослых мыши и кролика. Ответвление в виде зигзага показывает период времени, в течение которого дублированный ген после утраты функции очень быстро накапливал мутации.

три функциональных α -глобиновых гена расположены в виде кластера в хромосоме 11. Более того, по обеим сторонам $\alpha\psi 3$ -гена имеются ретровирусоподобные элементы, т. е. можно предположить, что определенную роль при образовании этого псевдогена сыграли ретровирусы).

Для реконструкции эволюционной истории псевдо- α -гена Мията и Ясунага построили молекулярно-филогенетическое древо, включающее три гена (рис. 7.11). В ходе своего анализа Мията и Ясунага [364] обратили внимание, что третий экзон мышинного псевдо- α -глобинового гена отличается от других экзонов крайней консервативностью. Это говорит о вероятном происхождении данной части псевдогена $M\psi\alpha$ от гена Ma или от одного из родственных α -глобиновых генов в результате двойного неравного кроссинговера (или генной конверсии) в относительно недавнее время, поэтому авторы исключили данный участок из рассмотрения. Они предположили также, что дупликация гена, приведшая к образованию $M\psi\alpha$, произошла в линии мыши T лет назад и что до момента t дублированный ген эволюционировал с той же скоростью, что и нормальный α -глобиновый ген. Обозначим скорость несинонимичных замен через k_A , а синонимичных — через k_S . Вероятно, по истечении какого-то

времени псевдоген $M\psi\alpha$ стал неактивным и все сайты в нем эволюционировали дальше с одинаковой скоростью, которую мы обозначим k_0 .

Предположив, что дивергенция кролика и мыши (T_0 на рис. 7.11) произошла 75 млн. лет назад, авторы получили оценку $T=24$ млн. лет для времени, прошедшего с момента дупликации, и $t=17$ млн. лет для времени с момента потери функции. Они вычислили также, что $k_A=0,82 \cdot 10^{-9}$, $k_S=6,6 \cdot 10^{-9}$ и $k_0=12,6 \cdot 10^{-9}$ на сайт за год, а $k_0/k_c=1,91$. Следовательно, скорость мутационных замен в гене $M\psi\alpha$ в 1,9 раза выше, чем скорость синонимичных замен в функциональных α -генах. Это означает, что в нормально функционирующем гене имеются какие-то селективные ограничения, в связи с чем не все синонимичные мутации оказываются селективно нейтральными. Мията и Хаясида [361] считают, что скорость эволюции этого псевдогена может соответствовать максимальной скорости, предсказываемой теорией нейтральности [236], т. е. отвечает $f_0=1$ в уравнении (5.1). Вопрос о том, действительно ли в этом случае скорость эволюции, выраженная через число мутационных замен, равна скорости точковых мутаций на гамету (если обе их измерять в годах), крайне важен для проверки состоятельности теории нейтральности. Это одна из проблем, которую еще предстоит решить.

Аналогичный анализ провели Такахата и Кимура [542], использовавшие модель более общего характера, чем модель 3ST, представленная на рис. 7.8. Мы последовали примеру Мияты и Ясунаги [364], исключив из анализа третий экзон. Главное отличие нашего анализа состоит в предположении, что ген $M\psi\alpha$ утратил функцию сразу же после своего возникновения путем генной дупликации (это соответствует $t=T$ на рис. 7.11). Мы рассчитали скорость эволюции отдельно для трех положений кодона, обозначив через k_i скорость для i -го положения ($i=1, 2$ или 3) в случае нормального гена. Предполагалось также, что дупликация произошла T_d лет назад и сразу же после этого дуплицированный ген стал «мертвым» и начал эволюционировать со скоростью k_i' вместо k_i . Для первого, второго и третьего положений кодона были получены следующие отношения: $k_1'/k_1=11,5$, $k_2'/k_2=13,9$ и $k_3'/k_3=0,9$. Отношения T_d/T_0 для трех кодоновых положений оказались равными 0,26; 0,42 и 0,43 со средним 0,37. Таким образом, если мы предполагаем, что дивергенция мышь — кролик произошла $T_0=8,0 \cdot 10^7$ лет назад, то $T_d=3,0 \cdot 10^7$ лет, т. е. «мертвый» ген появился около 30 млн. лет назад путем дупликации. Хотя эти величины и являются приблизительными, скорее всего скорости замен оснований в первых двух положениях кодона увеличились примерно в 10 раз после того, как ген утратил свою функцию, что привело к полному снятию селективных ограничений.

Изысканный статистический анализ скорости эволюции псевдогенов провели Ли и др. [312]. Они использовали данные по последовательностям псевдогена $\psi\alpha 1$ глубина человека [475], псевдогена $\psi\beta 2$ глубина кролика [287] и мышинного псевдогена $\psi\alpha 3$ [421]. Сравнив каждый из них с соответствующим нормальным аналогом и с менее близкими последовательностями таким же образом, как это было сделано Миятой и Ясунагой (см. рис. 7.11), Ли и др. оценили скорость нуклеотидных замен по первому, второму и третьему положениям кодона (в наших обозначениях k_1 , k_2 и k_3), время с момента дупликации (T на рис. 7.11) и время с момента утраты функции (t на рис. 7.11). Они оценили также скорость замен после утраты функции гена, предположив, что эта скорость (обозначим ее через k_0) одинакова для всех трех положений кодона. Для оценки t и k_0 был использован метод наименьших квадратов.

Полученные при этом значения усреднили для трех случаев (в группах по три) и получили следующий результат. Средние скорости замен для первого, второго и третьего положений кодонов в нормальных геновых генах равны $k_1 = 0,71 \cdot 10^{-9}$, $k_2 = 0,62 \cdot 10^{-9}$ и $k_3 = 2,64 \cdot 10^{-9}$, а скорость замен на сайт псевдогенов для «нефункционального» периода равна $k_0 = 4,6 \cdot 10^{-9}$ (за год). Авторы отмечают, что величина $4,6 \cdot 10^{-9}$ — одна из самых больших для скоростей нуклеотидных замен, известных до сих пор.

Хотя оценка k_0 , полученная Ли и др., составляет меньше половины соответствующей оценки Мияты и Ясунаги, отношение $k_0/k_3 = 1,74$ не сильно отличается от подобного отношения $k_0/k_s = 1,91$, полученного Миятой и Ясунагой. Таким образом, скорость синонимичных замен, даже будучи высокой, все же не достигает максимальной скорости, предсказываемой теорией нейтральности; следовательно, на синонимичные мутации все же действует отрицательный отбор, хотя его интенсивность может быть крайне малой. Эту проблему мы обсудим в следующем разделе в связи с «неслучайным» использованием синонимичных кодонов. Итак, различные данные об эволюции псевдогенов наиболее естественным образом можно объяснить в рамках теории нейтральной молекулярной эволюции.

7.5. «Неслучайное» использование синонимичных кодонов

Четко установлено, что синонимичные замены оснований (т. е. замены, которые не влекут за собой замещений аминокислот) происходят в ходе эволюции гораздо чаще, чем несинонимичные, например наиболее сильно такие различия выражены в случае гистона H4 (разд. 7.3). Этот факт очень хорошо согласуется с предсказанием теории нейтральности, так как по сравнению с несинонимичными заменами вероятность изменения фе-

нотипа данного организма в случае синонимичных замен неизмеримо меньше, поэтому при прочих равных условиях они должны гораздо меньше подвергаться дарвиновскому естественному отбору.

С другой стороны, становится все более очевидным, что синонимичные, или «вырожденные», кодоны используются во многих генах разных организмов совсем не одинаково, т. е. «неслучайным» образом (см., например, [144, 145]). По-видимому, неслучайное использование кодонов — скорее правило, чем исключение, и его часто упоминают как свидетельство против теории нейтральности. Здесь мы имеем такую же ситуацию, как и несколько лет назад, когда в качестве аргумента против теории нейтральности приводили данные о консервативной природе аминокислотных замен (см. разд. 7.1). Теперь мы знаем, что существование селективных ограничений никоим образом не противоречит теории нейтральности.

В этом разделе я попытаюсь показать, что и неслучайное использование синонимичных кодонов можно вполне удовлетворительно объяснить в рамках этой теории, хотя данная проблема более сложна и требует тщательного анализа.

Один из наиболее показательных примеров неслучайного использования синонимичных кодонов дают кодоны, соответствующие аминокислоте лейцину (Leu). Согласно стандартной таблице кода, лейцин кодируется шестью кодонами: UUA, UUG, CUU, CUC, CUA и CUG. Среди них бактериями используется главным образом кодон CUG [146]. Это видно из рис. 7.12, где под рубрикой «ЕСО» перечислены относительные частоты ($\times 10^3$) использования 61 кодона (исключая иницирующий и терминирующий кодоны) по 25 генам *E. coli*. Частота использования кодона CUG составляет около 60% от соответствующей величины для всех шести лейциновых кодонов. Совершенно другая картина наблюдается в случае дрожжей (*Saccharomyces*); относительные частоты использования кодонов для этого организма приведены в следующей графе рис. 7.12 под рубрикой «SAC». Так, например, среди шести лейциновых кодонов UUG используется гораздо чаще других (примерно в 80% случаев). В последней графе под рубрикой «НОМ» перечислены частоты использования различных кодонов для человека. Как следует из всех этих данных, по характеру использования кодонов человек ближе к бактериям, чем к дрожжам.

Наиболее детальный на сегодняшний день анализ использования кодонов выполнен Грэнсемом и др. [145—147]. Эти авторы составили обширные таблицы использования кодонов по известным последовательностям мРНК разных организмов. Число таких последовательностей насчитывало 90 в январе 1980 г., 119 в июне 1980 г. и 161 в 1981 г. Грэнсем и др. применили так называемый «анализ соответствий» (correspondence

analysis — метод многомерной статистики) к расстояниям между мРНК, найденным на основании различий в использовании 61 кодона. Каждая мРНК представлялась точкой в многомер-

ECO SACHOM	ECO SACHOM	ECO SACHOM	ECO SACHOM
Phe UUU 18 6 22 UUC 18 30 38 Leu UUA 7 3 1 UUG 7 50 4	Ser UCU 17 36 11 UCC 13 21 28 UCA 8 3 9 UCG 9 3 10	Tyr UAU 14 6 24 UAC 12 29 18 Term UAA — — — UAG — — —	Cys UGU 5 14 12 UGC 6 3 26 Term UGA — — — Trp UGG 12 6 8
Leu CUU 8 3 5 CUC 8 0 30 CUA 2 6 14 CUG 47 0 64	Pro CCU 5 3 10 CCC 4 0 12 CCA 7 30 16 CCG 19 0 5	His CAU 16 6 12 CAC 9 17 24 Gln CAA 13 14 10 CAG 29 0 38	Arg CGU 30 0 1 CGC 21 0 11 CGA 3 0 0 CGG 4 0 5
Ile AUU 24 29 5 AUC 32 28 17 AUA 5 0 2 Met AUG 22 18 20	Thr ACU 21 27 11 ACC 22 35 20 ACA 5 6 17 ACG 10 0 12	Asn AAU 10 6 4 AAC 25 38 38 Lys AAA 46 20 17 AAG 18 79 41	Ser AGU 7 0 10 AGC 9 0 12 Arg AGA 5 34 4 AGG 2 0 20
Val GUU 28 44 6 GUC 9 43 18 GUA 21 0 4 GUG 17 6 40	Ala GCU 37 54 20 GCC 20 33 26 GCA 31 0 9 GCG 23 0 2	Asp GAU 25 15 20 GAC 27 39 32 Glu GAA 37 47 21 GAG 18 8 34	Gly GGU 33 84 12 GGC 29 9 33 GGA 4 0 0 GGG 6 6 12

Рис. 7.12. Частоты использования кодонов ($\times 10^3$) для трех организмов — бактерий, дрожжей и человека (исключая иницирующий и терминирующий кодоны). Под рубрикой «ЕСО» представлены наблюдаемые частоты кодонов для 25 генов *E. coli* (данные рис. 3(b) из работы [146]), под рубрикой «SAC» — средние частоты кодонов для трех генов, кодирующих алкогольдегидрогеназу 1, изо-1-цитохром *c* и глицеральдегид-3- PO_4 -дегидрогеназу из *Saccharomyces cerevisiae* (по данным обзора [145]), и, наконец, под рубрикой «НОМ» приведены аналогичные частоты для четырех генов человека, кодирующих альфа-хорионический гонадотропин, β -глобин, пресоматотропин и хорионический соматомаммотропин (по данным работы [145]).

ном пространстве. Затем эти данные проецировались на плоскость, где горизонтальная и вертикальная оси соответствовали первому и второму по значимости факторам. Группирование проводили путем автоматической классификации. Результаты анализа 119 мРНК [147] показали, что большинство генов данного генома имеют одинаковую «кодирующую стратегию» в отношении выбора среди синонимичных кодонов. Таким образом,

гены млекопитающих, бактерий, вирусов, митохондрий, дрожжей и миксомицетов попадают в разные классы. Другими словами, существует некая согласованность в выборе вырожденных оснований, что подтверждает «геномную гипотезу» Грэнсема и др. [147], в рамках которой каждый ген того или иного генома стремится придерживаться характерного для своего вида принципа использования каталога кодонов.

С другой стороны, применение «анализа соответствий» к частотам встречаемости аминокислот в 119 белках не выявило никакого группирования белков по типу генома: белки вирусов, бактерий, млекопитающих и других организмов попадают в один класс. Это означает, что в мРНК содержится не только та информация, которая необходима для кодирования белков, но и «избыточная» информация, которая связана главным образом с выбором вырожденных оснований.

Опираясь на эти наблюдения, Грэнсем [144] утверждает, что мРНК представляет собой самостоятельную эволюционную структуру, и считает, что им и его сотрудниками обнаружена независимая от белка молекулярная эволюция «ненейтрального характера». Однако, по-моему, то, что он назвал «ненейтральным», является скорее проявлением селективного ограничения, нежели адаптивной мутационной замены, которую, очевидно, имел в виду Грэнсем.

Проблема, следовательно, состоит в том, чтобы установить главную причину селективного ограничения, которая обуславливает определенный выбор вырожденных кодонов, характерный для каждого геномного типа. Значительный шаг на пути к выяснению этой причины был недавно сделан Икемурой [189—191]. С помощью двумерного электрофореза в полиакриламидном геле он разделил 26 известных тРНК *E. coli* и определил их относительные количества, измеряемые числом молекул в клетке. Затем он исследовал зависимость между содержанием различных тРНК и частотой использования соответствующих кодонов в генах *E. coli*, таких как гены триптофансинтетазы А (*trp* А), репрессора *lac*-оперона (*lacI*) и рибосомного белка (г-белка). Им обнаружена очень сильная корреляция между содержанием соответствующих тРНК и использованием синонимичных кодонов, в том числе и для кодонов, отвечающих различным аминокислотам. Для нас особенно интересен тот факт, что среди синонимичных кодонов для данной аминокислоты наиболее часто используемый кодон однозначно соответствует наиболее многочисленному виду изоакцептирующих тРНК. Эту зависимость иллюстрирует рис. 7.13 для шести синонимичных кодонов, кодирующих лейцин. Черные столбики отвечают частотам использования кодонов, а заштрихованные — содержанию соответствующих тРНК. Четкая корреляция этих величин очевидна. Отметим, что кодоны CUC

и CUU узнаются одной и той же тРНК ($\text{тРНК}_2^{\text{Leu}}$); то же самое относится и к кодонам UUG и UUA.

Свой вывод о том, что частота использования кодона определяется в основном наличием соответствующей тРНК, Икемура подтвердил следующим открытием: такая же корреляция обнаружилась и у дрожжей, где характер использования кодо-

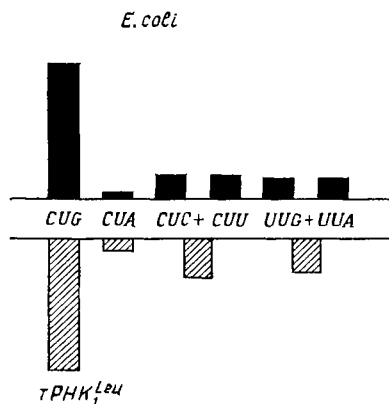


Рис. 7.13. Коррелиция между относительными частотами использования кодонов лейцина (черные столбики) и относительным содержанием соответствующих тРНК (заштрихованные столбики) у *E. coli*. Данные о частотах использования кодонов взяты из рис. 7.12, а о содержании тРНК — из работ [190, 191]. Знак «+» между кодонами CUC и CUU означает, что эти кодоны узнаются одной и той же тРНК ($\text{тРНК}_2^{\text{Leu}}$), содержание которой (если принять содержание $\text{тРНК}_2^{\text{Leu}}$ за единицу) представлено заштрихованным столбиком под символами «CUC+CUU». Аналогичный смысл имеет и знак «+» между UUG и UUA. Содержание тРНК, соответствующей кодону CUA, который используется реже всех остальных, равно 0,1 (личное сообщение Икемуры).

нов существенно отличается от такового у *E. coli*. Например, среди шести синонимичных кодонов, кодирующих лейцин, у дрожжей чаще используется UUG. Икемура обнаружил (личное сообщение, июль 1981), что в этом случае самым многочисленным типом тРНК является как раз тот, который узнает UUG (рис. 7.14). Эти открытия заставляют сомневаться в ценности предположений, высказанных Модино и др. [367], о том, что UUA и UUG являются «предтерминирующими кодонами» (которые могут мутировать с образованием терминирующих кодонов в результате одной нуклеотидной замены) и что, поскольку лейцин кодируется еще и четырьмя другими, «непредтерминирующими кодонами», UUA и UUG никогда не используются в нормальных генах α - и β -цепей глобинов. Авторы упомянутой работы утверждают, что организмы стремятся по мере возможности не использовать «предтерминирующие кодоны» при выборе эволюционной стратегии, направленной на уменьшение

частоты мутаций с сильно выраженными эффектами. Мне кажется, что получаемое при этом селективное преимущество слишком мало (вероятно, не более скорости мутирования), чтобы такая стратегия имела эволюционные последствия.

Несколько ранее Фитч [115] проанализировал последовательности мРНК β -цепей гемоглобинов человека, мыши и кролика и выявил значительное отклонение в использовании кодонов, отличающихся от терминирующих кодонов только од-

Дрожжи

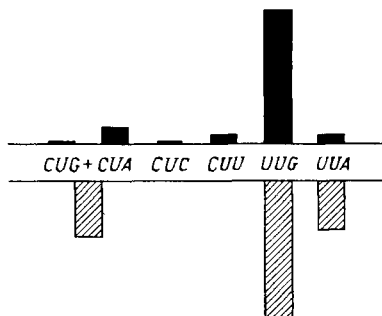


Рис. 7.14. Корреляция между частотами использования синонимичных кодонов лейцина и относительным содержанием соответствующих тРНК у дрожжей). Как и на рис. 7.13, черными столбиками представлены относительные частоты использования кодонов, а заштрихованными — относительные количества тРНК. Частоты использования получены из данных рис. 7.12, а сведения об относительных количествах тРНК представлены Икемурой (личное сообщение).

ним основанием. Редкое использование кодонов UUA и UUG оказалось свойственным и другим генам человека (рис. 7.12); по-видимому, если гипотеза Икемуре верна, это обусловлено малочисленностью той тРНК, которая узнает соответствующие два кодона, а не тем, что они представляют собой «предтерминирующие кодоны».

Допустим, что выбор синонимичного кодона определяется в основном наличием соответствующей тРНК и что это, как считают Пост и др. [469] и Икемура [191], связано с эффективностью трансляции. Отсюда может следовать, что характер предпочтительного использования кодонов отражает то оптимальное состояние, при котором популяция синонимичных кодонов отвечает популяции соответствующих тРНК в клетке. Такое состояние должно обеспечивать более эффективное выполнение клеткой ее функций, а значит, и большую приспособленность по Дарвину. К анализу этой проблемы можно применить далее концепцию стабилизирующего отбора [245]. Такой подход должен соответствовать также и «геномной гипотезе» Грэнсема и др. [147], согласно которой в выборе «вырож-

денных» оснований между генами одинаковых или сходных геномов существует удивительная согласованность и что «единицей отбора является геном, а не отдельный ген».

Далее мы предполагаем, что характер использования кодонов в ходе эволюции изменяется таким образом, чтобы отвечать набору разных тРНК. Однако может иметь место и обратная ситуация: соотношение между разными тРНК (определяемое числом их генов) «подгоняется» к относительным частотам синонимичных кодонов. Какая из этих возможностей реализуется на самом деле, мы не знаем. Вполне вероятно, что такие изменения взаимосвязаны, т. е. одно из них регулирует другое, и осуществляются они очень медленно и согласованно. Однако здесь я придерживаюсь того же мнения, что и Икемура [190, 191], т. е. полагаю, что частота использования кодонов определяется соотношением между разными тРНК.

Вернемся к результатам, полученным в конце предыдущей главы (см. разд. 6.8), где разбирались вопросы вероятности фиксации мутантного аллеля и распределения аллельных частот (в случае двух аллелей). Рассмотрим отдельный сайт, соответствующий третьему положению в кодоне, и для простоты сгруппируем четыре основания в два класса, обозначив основания U (урацил) и A (аденин) через A_1 , а G (гуанин) и C (цитозин) — через A_2 . Обозначим частоту A_2 через x ; тогда частота A_1 будет равна $(1-x)$. Если v_1 — скорость мутирования A_1 в A_2 , а v_2 — скорость обратного процесса, то уравнение (6.65) можно использовать для описания устойчивого распределения x под давлением мутаций и случайного дрейфа.

Пусть $Q_{\text{опт}}$ — оптимальная доля A_2 (G или C) по третьему положению в кодоне. Предположим, что скорости мутирования между A_1 и A_2 равны, т. е. в уравнении (6.65) $V_1 = V_2$. Правомочность этого предположения мы обсудим позже. Поскольку ожидаемое при учете только мутационного процесса среднее значение для x (т. е. $\bar{x}$) не совпадает с $Q_{\text{опт}}$, если только $Q_{\text{опт}}$ не равно 0,5, то мы полагаем, что стабилизирующий отбор таков, что $\bar{x}$ удерживается вблизи $Q_{\text{опт}}$. При этом по индивидуальным сайтам аллель A_2 большую часть времени находится в состоянии фиксации либо элиминации (рис. 6.8). Пусть f_2 — вероятность того, что A_2 фиксирован (а A_1 элиминирован). Аналогичным образом пусть f_1 — вероятность того, что A_1 фиксирован (а A_2 элиминирован). Тогда из соотношения [6.68] мы имеем $f_2/f_1 = \exp(B_1)$, где $B_1 = 4N_e\beta_1 = -4N_e\lambda ta$. Таким образом, $B_1 = \ln(f_2/f_1)$ и при $f_2/f_1 = 0,7/0,3$ мы получим $B_1 = 0,85$. Так как $-\lambda ta = B_1/(4N_e)$, а N_e для большинства видов млекопитающих должно быть не меньше 10^4 , то интенсивность отбора, который действует на индивидуальный сайт, формируя «неслучайное» использование кодонов, крайне мала, что открыва-

ет широкое поле деятельности для случайного дрейфа. Это соответствует высказыванию Лэттера [294] о том, что мутации, ответственные за полиморфизм ферментов, имеют очень слабый вредный эффект, так что значения Ns не выходят за пределы 1—3.

Остается еще один важный вопрос: как сильно влияет такой отбор на скорость эволюции, выражаемую через число мутационных замен? Поскольку скорость эволюции при стабилизирующем отборе, отнесенная к скорости эволюции для строго «нейтральной» ситуации, равна u/u_0 , где $u_0 = 1/(2N)$, то из уравнения (6.60) мы имеем

$$u/u_0 = 1 \int_0^1 \exp[-B_1x + B_2x(1-x)] dx, \quad (7.8)$$

где $B_1 = -4N_e\lambda ma$, $B_2 = 2N_e\lambda(1-\lambda m^2)a^2$. Если предположить, что $2N_e\lambda a^2$ пренебрежимо мало, так что $B_2 \approx 0$, то мы получим

$$u/u_0 \approx 2f_1f_2 \ln(f_2/f_1)/(f_2 - f_1). \quad (7.9)$$

При $f_2/f_1 = 0,7/0,3$ (ситуация, наблюдаемая для третьего положения кодона в глобиновой и в других мРНК млекопитающих) имеем $u/u_0 = 0,89$. Иными словами, молекулярная эволюция идет медленнее примерно на 10% по сравнению с ситуацией, ожидаемой для условий полной селективной нейтральности. А в случае, когда $f_2/f_1 = 0,9/0,1$, $u/u_0 \approx 0,49$, что означает почти 50%-ное замедление темпов замен. Если B_2 не пренебрежимо мало, степень замедления увеличивается. При $f_2/f_1 = 0,7/0,3$ получим $u/u_0 = 0,63$, если $B_2 = 2$. Это соответствует случаю распределения частоты гена, показанному на рис. 6.8 сплошной кривой. При $B_2 = 4$ отношение $u/u_0 = 0,44$, а если $B_2 = 10$, оно составляет 0,14; вероятно, более реалистичны значения $B_2 < 2$, так как a скорее всего мало и, следовательно, a^2 пренебрежимо мало.

Однако на самом деле ситуация является более сложной. Во-первых, число возможных «аллелей» (оснований) по каждому нуклеотидному сайту равно четырем, а не двум; неравенство частот U и A, а также G и C может еще больше замедлять скорость эволюции, хотя на величине u/u_0 это может мало сказаться. Во-вторых, степень селективного ограничения может быть неодинаковой для разных генов и зависеть от их активности. Согласно Икемуре [190, 191], степень изменения приспособленности ($\Delta\omega$), обусловленного изменением активности гена (Δr) вследствие синонимичной замены, задается уравнением

$$\Delta\omega = Pn \cdot \Delta r, \quad (7.10)$$

где Pn — параметр, тесно связанный с активностью гена [191]. Это означает, что для данного изменения (Δr), вызванного си-

нонимичной заменой, степень изменения приспособленности (Δw) пропорциональна Pn . Иначе говоря, более сильное влияние на приспособленность оказывают гены преобладающих белков (такие, как гены г-белков). Икемура считает, что доля мутаций, которые можно рассматривать как нейтральные или почти нейтральные, должна быть больше для генов с низкими значениями Pn , чем с высокими.

Уравнение Икемуры (7.10) можно представить в более общем виде, соответствующем теории нейтральности:

$$\bar{s}' \propto Pn \cdot |a|, \quad (7.11)$$

где $\bar{s}'$ — средний коэффициент отбора против данного мутантного аллеля, $|a|$ — эффект замещения аллеля данной мутантной формой.

Что касается различий в содержании разных тРНК, то Икемура [190] показал, что у *E. coli* их количество пропорционально числу соответствующих генов; наиболее многочисленные тРНК кодируются несколькими копиями гена, а самые немногочисленные — одной копией.

С другой стороны, обоснованность нашего предположения о том, что скорости мутирования (это не обязательно скорости эволюционных замен оснований) одинаковы для всех четырех оснований (а следовательно, и для A_1 и A_2), по-видимому, хорошо подтверждается данными, полученными при сравнении нуклеотидного состава псевдогенов и соответствующих нормальных генов. Исследуя скорость эволюции псевдогенов, Мията и Хаясида [361] сделали интересное наблюдение: кодоны, оканчивающиеся на U, C, A или G, в глобиновых псевдогенах встречаются с более близкими частотами, чем в их функционирующих партнерах. Тем не менее даже в псевдогенах, по-видимому, в значительной мере сохраняется тот нуклеотидный состав, который был характерен для них в период их функционирования. Я рассматривал только частоты оснований по третьим положениям кодонов, поскольку именно по ним глобиновые гены и псевдогены отличаются друг от друга.

Если псевдогены свободны от селективных ограничений и замены оснований происходят у них с большой скоростью, то состав оснований по набору третьих положений кодонов, выбранному из псевдогенов, должен лучше отражать характер мутационных изменений, не подвергающихся действию отбора. В табл. 7.5 приведены частоты четырех оснований по третьему положению кодонов четырех глобиновых псевдогенов. Эти частоты были получены при сравнении $\alpha 3$ - и $\alpha 1$ -генов мыши [421], $\psi\beta 2$ - и $\beta 1$ -генов кролика [237], $\psi\alpha 1$ - и $\alpha 2$ -генов человека [475] и $\beta h 3$ - и β_{maj} -генов мыши [361]. Из этой таблицы ясно, что все четыре основания встречаются примерно с одинаковыми частотами (подтверждая тем самым предположение, что

Таблица 7.5. Наблюдаемые числа оснований, находящихся в третьих положениях кодонов глобиновых псевдогенов. При подсчете частот оснований учитывались только те третьи положения, по которым данный псевдоген отличается от одного из своих функционирующих аналогов у того же самого вида

Псевдоген	Наблюдаемые числа				Всего
	U	C	A	G	
$\alpha 3$ -ген мыши	11	6	6	6	29
$\psi\beta 2$ -ген кролика	9	11	9	5	34
$\psi\alpha 1$ -ген человека	17	7	10	10	44
$\beta h 3$ -ген мыши	1	7	11	6	25
Всего	38	31	36	27	132
%	28,8	23,5	27,3	20,4	100

$V_1 = V_2$). С другой стороны, частоты оснований по третьим положениям кодонов в нормальных аналогах этих псевдогенов существенно различаются. На рис. 7.15 сравниваются частоты оснований (в %), полученные таким образом, для глобиновых псевдогенов (черные столбики) и их функционирующих аналогов (заштрихованные столбики). Интересно отметить, что последние очень сходны не только с нуклеотидным составом по

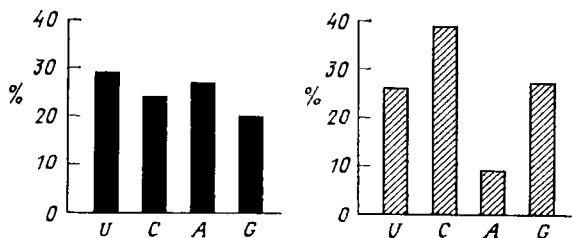


Рис. 7.15. Относительные частоты четырех оснований по третьим положениям кодонов для глобиновых псевдогенов (слева) и для функционирующих глобиновых генов (заштрихованные столбики). Подробности см. в тексте.

третьему положению кодонов глобиновых последовательностей вообще, но и с нуклеотидным составом по третьему положению «квартетов» синонимичных кодонов в генах млекопитающих (см. работу [144], рис. 3). Мията и Хаясида [361], по-видимому, правы, утверждая, что скорость эволюции псевдогенов — это некая предельная скорость, ожидаемая в рамках теории нейтральности [236], и предполагая, что на синонимичные изменения все же наложены некоторые селективные ограничения.

Пытаясь оценить долю всех точковых мутаций, которые оказались эволюционно неприемлемыми, Бойер и др. [28] пришли, в частности, к следующим выводам. 1. По крайней мере 95% несинонимичных замен оказываются неприемлемыми для β -гемоглобиновых генов у приматов. 2. Примерно 70% синонимичных мутаций β -Нб адаптивно неприемлемы. 3. Лишь около 11% всех видов точковых мутаций оказываются допустимыми, и скорость мутирования составляет по крайней мере 10^{-8} на сайт за год. 4. Неслучайность в выборе кодонов возрастает благодаря отбору, действовавшему на претрансляционные процессы. 5. Селективное ограничение, налагаемое на нуклеотидные дубликаты, не только влияет на несинонимичные изменения в эволюции, но и вносит вклад в происхождение систем генетического полиморфизма.

Хотя эти выводы требуют некоторого пересмотра в свете более поздних исследований, которые я уже обсудил, Бойер и др. [28] оказались правы в том, что не все синонимичные мутации строго нейтральны и в действительности на значительную их часть действует отрицательный отбор (проблема дефицита CpG-дублетов обсуждается в работе [20] и была затронута нами в конце гл. 3).

По-видимому, самый замечательный факт, обнаруженный в ходе различных исследований, суммированных в этом разделе, состоит в том, что из-за неслучайности выбора синонимичных кодонов эволюция (на уровне замен оснований) замедляется, а не ускоряется. Он подтверждает положение теории нейтральности о том, что действующий при этом отбор является отрицательным (и является следствием стабилизирующего отбора), а не положительным дарвиновским отбором, о котором обычно говорят сторонники селекционистской теории.

Хотя многое здесь еще предстоит выяснить, я удовлетворен тем, что такие разные факты, как быстрое изменение псевдогенов и неслучайное использование синонимичных кодонов, можно объяснить — и притом на количественном уровне — в рамках теории нейтральности с привлечением концепции стабилизирующего отбора. Но и в этом случае остается очевидным первостепенное значение случайного генетического дрейфа для эволюции на молекулярном уровне.

Молекулярная популяционная генетика

8.1. Для чего необходим стохастический подход

Эта глава служит как бы введением к следующей, где обсуждается механизм поддержания генетической изменчивости на молекулярном уровне. Поскольку в рамках теории нейтральности внутривидовая изменчивость рассматривается с позиций стохастической теории молекулярной эволюции, необходимо иметь четкое представление о теории случайных процессов, в особенности применительно к динамике молекулярных мутаций в конечных популяциях. Вообще говоря, я уже касался теории случайных процессов в гл. 3, а также в разд. 6.8. В настоящей главе я более последовательно изложу соответствующие аспекты этой теории, рассматривая изменение генных частот как стохастический процесс. Термин «стохастические процессы» используется здесь применительно к случайным событиям, происходящим во времени.

Необходимость применения стохастического подхода обусловлена уникальностью мутаций и конечностью размера популяций; отдельные молекулярные мутации благодаря их крайне низкой частоте могут считаться практически уникальными, поэтому в конечной популяции судьба таких мутаций в основном определяется случайной выборкой гамет при размножении.

Для того чтобы проиллюстрировать все эти положения на конкретном примере, рассмотрим популяцию некоего гипотетического млекопитающего. Пусть продолжительность поколения в этой популяции равна одному году, а ее размер ежегодно составляет 100 000 особей. Для простоты проанализируем пока идеальный случай, когда эффективный размер популяции равен ее действительному размеру ($N_e = N = 10^5$). Рассмотрим генный локус, кодирующий условный белок, который состоит из 333 аминокислот (т. е. близок по величине к алкогольдегидрогеназе). Такой белок соответствует 999 нуклеотидам кодирующего участка гена. Предположим также, что скорость мутирования на сайт на поколение, $\nu_{\text{нукл}}$, равна $5 \cdot 10^{-9}$ и что скорости мутирования для четырех оснований (А, С, G и Т) одинаковы, т. е. мутационная замена одного основания на другое

(скажем, С на А) происходит с частотой $(5/3) \cdot 10^{-9}$. Отсюда ожидаемое число новых мутаций на сайт на поколение в популяции равно $2Nv_{\text{нукл}}$, или 0,001. При столь низкой скорости мутирования вновь появившийся мутантный ген, модифицированный по определенному сайту, скорее всего будет представлен в популяции в единственном числе. При этом мутации по данному сайту будут появляться в популяции в среднем с интервалом в 1000 поколений. Более того, можно показать, что подавляющее большинство мутаций случайным образом утратится популяцией в течение примерно 20 поколений.

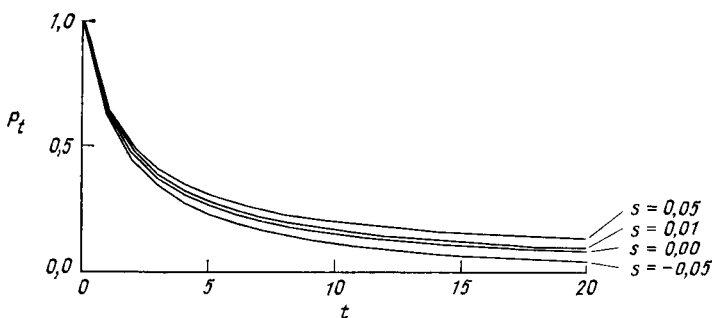


Рис. 8.1. Вероятность сохранения мутации как функция времени (в поколениях), прошедшего после ее возникновения. s — коэффициент отбора мутаций. При расчетах использовалось уравнение (8.1).

Скорость мутирования $v_{\text{нукл}} = 5 \cdot 10^{-9}$ получена из оценки скорости эволюции глобиновых псевдогенов, которая в рамках теории нейтральности равна скорости мутирования (разд. 7.4). Величину скорости мутирования можно получить другим путем, не требующим предположения о нейтральности и основанном на данных о частоте и распределении редких вариантов гемоглобина человека среди населения Японии. Используя эти данные, Кимура и Ота [259] определили скорость мутирования на аминокислоту $v_{aa} = 4,5 \cdot 10^{-9}$, что соответствует $v_{\text{нукл}} = 1,5 \cdot 10^{-9}$. Принимая во внимание приближенный характер этих оценок, можно считать, что соответствие между ними достаточно хорошее, и мы можем считать величину $5 \cdot 10^{-9}$ репрезентативной.

На рис. 8.1 представлена вероятность сохранения вновь появившейся мутации, вычисленная по рекуррентной формуле

$$P_t = 1 - \exp \{ -(1+s) P_{t-1} \}, \quad (8.1)$$

для нескольких случаев с разными коэффициентами отбора. Здесь P_t — вероятность сохранения мутации через t поколений ($t=0, 1, 2, \dots$) при $P_0=1$. Эта задача была впервые решена Фишером [106, 107] (объяснение см. в работе [67], с. 421).

Необходимо отметить следующий факт, который, несмотря на огромное влияние работ Холдейна и Фишера, все же иногда упускается из виду: обычная судьба отдельных мутаций — их исчезновение из популяции. Это относится даже к мутациям, имеющим явное селективное преимущество ($4N_s \gg 1$), если оно, конечно, не слишком велико.

Ниже я остановлюсь на поведении нейтральных мутаций более подробно, поскольку их важность для данной книги совершенно очевидна. Следует также учитывать, что поведение мутации, подверженной очень слабому давлению отбора, во многом сходно с поведением нейтральной мутации. Среднее время фиксации селективно нейтральной мутации (исключая случаи ее элиминации) составляют $4N_e$ поколений [253], что для нашего примера равно 400 000. С другой стороны, среднее число поколений до вытеснения мутации (исключая случаи фиксации) равно

$$\bar{t}_0 = 2 (N_e/N) \ln 2N. \quad (8.2)$$

В нашем примере это составляет 24,4 поколения. Таким образом, существует огромная разница между средним временем фиксации мутации и средним временем ее элиминации. Стандартное отклонение для времени элиминации равно

$$\sigma(t_0) \approx 4N_e/\sqrt{N}, \quad (8.3)$$

что составляет примерно 1265 поколений и значительно больше среднего. С другой стороны, стандартное отклонение для времени фиксации равно $2,15 N_e$ поколений, т. е. составляет примерно половину среднего.

Хотя подавляющее большинство мутантных генов элиминируется, их меньшая, более «удачливая», часть распространяется по всей популяции. Вероятность фиксации нейтральной мутации равна $1/2 N$, т. е. для нашего примера 1 : 200 000.

В течение длительного времени эволюции происходит взаимозамещение аллелей: наиболее распространенное основание, занимающее какой-либо сайт, заменяется другим (например, С заменяется на А). В рассматриваемом нами случае среднее время, разделяющее два последовательно фиксированных аллельных состояния, примерно равно $1/v_{\text{нукл}}$, или $2 \cdot 10^8$ поколений.

Чтобы лучше понять механизм этого процесса, рассмотрим стационарное распределение генов. Поскольку возможное число различных оснований, занимающих данный сайт, равно четырем, для описания аллельных частот потребуются три параметра, а для графического представления — четырехмерное пространство. Однако, если скорости мутирования четырех оснований одинаковы, а мутации селективно нейтральны (т. е. эквивалентны), наш подход можно упростить, сосредоточив внима-

ние на одном из оснований (скажем, А) и сгруппировав остальные три основания в один класс (обозначим его А'). Пусть x — частота встречаемости основания А в данном сайте в популяции. Тогда стационарное распределение x задается формулой Райта

$$\phi(x) = \frac{\Gamma(\alpha + \beta)}{\Gamma(\alpha)\Gamma(\beta)} (1-x)^{\alpha-1} x^{\beta-1}. \quad (8.4)$$

В данном случае $\alpha = 4N_e v_{\text{нукл}}$, $\beta = \alpha/3$, а $\Gamma(\cdot)$ — гамма-функция (см. [230]; формула (8.4) — частный случай приведенного в

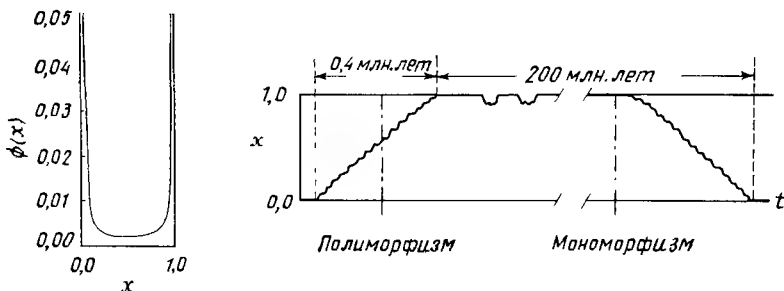


Рис. 8.2. Иллюстрация замещения оснований в данном нуклеотидном сайте у гипотетического млекопитающего, популяция которого состоит из 10^5 размножающихся особей, а продолжительность поколения равна одному году. Предполагается, что скорость мутирования на сайт на поколение $v_{\text{нукл}} = 5 \cdot 10^{-9}$ и одинакова для всех четырех нуклеотидов (А, С, G, Т). Слева представлено частотное распределение для основания А, справа — типичный ход изменения частоты А во времени. Подробности см. в тексте.

этой работе соотношения (3.5) для $K=4$. $\phi(x)dx$ можно рассматривать как вероятность того, что частота А лежит в малом интервале частот от x до $x+dx$. Это распределение допускает и другую, более приемлемую для данного контекста интерпретацию: если следить за данным сайтом в течение долгого (с эволюционной точки зрения) времени, то относительное число случаев, когда частота основания А равна x , будет пропорционально $\phi(x)$. На рис. 8.2 слева представлено распределение $\phi(x)$ для $N_e = N = 10^5$ и $v_{\text{нукл}} = 5 \cdot 10^{-9}$ ($\alpha = 0,002$). Это распределение имеет четко выраженную U-образную форму.

Вероятность временной фиксации в популяции основания А определяется путем интегрирования $\phi(x)$ от $x=1-1/(2N)$ до $x=1$, что дает

$$f_1 = C_1 (2N)^{-\alpha}, \quad (8.5)$$

где $C_1 = \Gamma(\alpha + \beta) / [\Gamma(\alpha + 1)\Gamma(\beta)]$ [230]. В нашем примере эта величина составляет 0,24397. Вероятность окончательной фиксации одного из четырех оснований равна $4f_1$, или 0,97588; тогда вероятность того, что в популяции одновременно присутст-

вуют 2 или больше разных оснований, равна $(1-4f_1)$, или 0,02412. В результате этих вычислений мы получаем вероятности достижения состояния фиксации, или мономорфизма (а также вероятности достижения состояния сегрегации, или полиморфизма), в строгом смысле слова. На практике, однако, при определении мономорфизма удобнее включать в этот класс и случаи, когда в популяции присутствуют редкие варианты, поскольку последние вряд ли будут обнаружены при ограниченном размере выборки.

Итак, назовем популяцию мономорфной, если суммарная частота аллельных вариантов (или редко встречающихся оснований) не превышает q (обычно q принимают равным 0,01). Согласно этому определению, вероятность того, что популяция мономорфна (на уровне q) по основанию А, равна

$$P_{\text{моно}}(A) = \int_{1-q}^1 \phi(x) dx \approx C_1 q^\alpha. \quad (8.6)$$

В нашем примере, где $\alpha=0,002$, $C_1 \approx 0,25$ и $q=0,01$, $P_{\text{моно}}(A) \approx 0,2477$. Это означает, что вероятность полиморфизма, определенная как $1-4P_{\text{моно}}(A)$, равна 0,009168.

Средняя гомозиготность по этому сайту равна

$$\bar{h}_0 = 4 \int_0^1 x^2 \phi(x) dx = (\alpha+3)/(4\alpha+3), \quad (8.7)$$

и для нашего случая составляет 0,998005. Тогда средняя гетерозиготность ($\bar{h}_e$) по этому сайту, $1-\bar{h}_0$, равна 0,001995.

Результаты этих вычислений свидетельствуют о том, что состояние полиморфизма по данному нуклеотидному сайту встречается крайне редко; это и обеспечивает существование видоспецифичных последовательностей нуклеотидов. У нашего гипотетического млекопитающего замена одного фиксированного состояния на другое происходит в среднем один раз в 200 млн. лет (поколений), тогда как переходное (полиморфное) состояние длится в среднем 400 000 лет, что составляет всего 1/500 от времени пребывания в состоянии фиксации. Ход таких изменений представлен на рис. 8.2 справа.

Перейдем теперь к рассмотрению эволюционных изменений на уровне целого гена. В нашем примере кодирующий участок гена представлен 999 нуклеотидными сайтами, что соответствует 333 кодонам. Отсюда суммарная скорость мутирования составляет около $5 \cdot 10^{-6}$ на ген на поколение. На этом участке все замены первого и второго нуклеотидов в кодонах, как правило, приводят к замене соответствующей аминокислоты, тогда как в случае третьего нуклеотида к такому эффекту приводит

только $1/3$ замен. Таким образом, суммарная скорость появления мутаций, приводящих к замене аминокислоты, равна примерно $3,9 \cdot 10^{-6}$. Обозначим эту скорость $v_{T(A)}$. Предположим [508], что в среднем $1/3$ замен аминокислот может быть обнаружена с помощью электрофореза. Тогда скорость возникновения мутаций, выявляемых электрофоретическими методами, равна $1,3 \cdot 10^{-6}$ на поколение. Обозначим эту скорость $v_{T(Э)}$. Большую часть этих изменений составят скорее всего вредные мутации, поэтому предположим, что в среднем только $1/10$ мутаций, сопровождающихся заменой аминокислот, селективно нейтральны, тогда как оставшиеся $9/10$ безусловно вредны и не вносят вклада ни в молекулярную эволюцию, ни в полиморфизм. При таких упрощающих, но все же не столь уж далеких от реальности предположениях скорость мутирования для выявляемых методом электрофореза нейтральных мутаций равна $1,3 \cdot 10^{-7}$ на ген на поколение. Обозначим эту скорость $v_0(Э)$.

Каким может быть уровень генетической изменчивости по данному локусу, проявляющейся в виде полиморфизма белков (выявляемого стандартными методами электрофореза)? Чтобы ответить на этот вопрос, воспользуемся «моделью бесконечного числа аллелей», которая будет более подробно рассмотрена в следующем разделе. Эта модель основана на допущении, что ген состоит из очень большого числа нуклеотидных сайтов, каждый из которых может быть занят одним из четырех оснований. Возможное число аллельных состояний такого гена астрономически велико. Для нашего случая это равносильно тому, что всякий раз, когда в белке происходит аминокислотная замена, появляется новый тип белка, не существовавший ранее в популяции. В рамках этой модели ожидаемая гетерозиготность по данному локусу равна

$$\bar{H}_e = \frac{4N_e v_0(Э)}{4N_e v_0(Э) + 1} \quad (8.8)$$

[246]. Если $N_e = 10^5$, а $v_0(Э) = 1,3 \cdot 10^{-7}$, то $H_e \approx 0,049$, или 4,9%. Для позвоночных, изученных разными авторами к настоящему времени, распределение средней гетерозиготности имеет моду в точке $\sim 5\%$ (см. гл. 9), так что наша оценка достаточно реалистична. Модель бесконечного числа аллелей позволяет также вычислить вероятность полиморфизма белков:

$$P_{\text{поли}} = 1 - q^V \quad (8.9)$$

[257], где $V = 4N_e v_0(Э)$. Для $q = 0,01$ получаем $P_{\text{поли}} = 0,213$, т. е., если исследовать большое число эквивалентных локусов в геноме нашего гипотетического млекопитающего, ожидаемая доля полиморфных локусов будет равна 21,3%. Это также вполне реалистичная оценка. Согласно Нево [420], среднее число полиморфных локусов на популяцию (P) и средняя ге-

гетерозиготность на локус на особь (H) равны соответственно $P=0,147\pm0,098$ и $H=0,0359\pm0,0245$ для 46 видов млекопитающих.

Возникает вопрос: какова скорость эволюции, выраженная через число замен аминокислот, у нашего гипотетического млекопитающего? Эта скорость обычно определяется путем сравнительного изучения гомологичных последовательностей в белках, причем за единицу времени принимается один год. В данной книге мы обозначили ее через k_{aa} (гл. 4). Теоретически эта скорость соответствует скорости возникновения нейтральных замен на гамету на поколение [см. уравнение (3.20)]. Мы предположили, что только $1/10$ мутаций, сопровождающихся заменой аминокислот, селективно нейтральны и что $v_{\text{нукл}}=5\cdot 10^{-9}$. Таким образом, соответствующая скорость мутирования на кодон (состоящий из трех нуклеотидов) равна $v_{0(A)}=(1+1+1/3) v_{\text{нукл}}$ ($1/10$) в год, что составляет $1,17\cdot 10^{-9}$. Отсюда для нашего гипотетического млекопитающего скорость молекулярной эволюции $k_{aa}=1,7\cdot 10^{-9}$ на аминокислотный сайт в год.

В гл. 4 мы уже видели, что медиана скоростей эволюции для 21 белка (определенная в основном для млекопитающих) равна $1,3\cdot 10^{-9}$ на аминокислотный сайт на поколение, так что полученная выше величина k_{aa} вполне реалистична. Отметим, что эта скорость не зависит от эффективного размера популяции, так как относится к селективно нейтральным заменам.

Наконец, я хотел бы подчеркнуть, что сделанное нами ранее упрощающее предположение о равенстве эффективной и фактической численности популяции совершенно не соответствует действительности. При рассмотрении эволюции и молекулярной изменчивости природных популяций эффективный размер популяции обычно значительно ниже, чем наблюдаемый. Если за время между моментом подсчета численности и временем размножения в популяции наблюдается высокая смертность или если в размножении принимают участие только несколько особей (как, например, самцы высокополигинного вида), эффективная численность популяции может оказаться на несколько порядков ниже численности, полученной путем подсчета особей. Как показано в гл. 3, к расхождению между этими величинами приводят такие факторы, как неравенство числа самцов и самок, принимающих участие в размножении, флуктуации размеров популяции от поколения к поколению, наследственное различие в плодовитости, вымирание отдельных колоний с последующей реколонизацией и т. д. Особенно ярко это проявляется в том случае, когда размер популяции в ходе эволюции периодически резко уменьшается (эффект «горлышка бутылки»), а последующее увеличение размера идет медленными темпами; в этих условиях средняя гетерозиготность сильно уменьшается и соответственно уменьшается величина N_e [401].

Следовательно, если, применяя уравнение (8.8), мы принимаем за N_e размер популяции данного вида в условиях его процветания, ожидаемое среднее значение гетерозиготности оказывается завышенным. Даже если какой-то вид в данное время широко распространен, занимает обширную территорию и состоит из огромного числа особей, это не означает, что ситуация была такой же миллионы лет назад и сохранится в последующие миллионы лет. Рано или поздно это состояние нарушится из-за процесса видообразования. Для «среднего» млекопитающего (такого, как в нашем примере) разумно предположить, что, если этот вид не вымирает, он превращается в другой приблизительно через каждые 500 тыс. лет, причем на каждом этапе видообразования происходит значительное уменьшение размера популяции. Вероятно, ни для одного вида на Земле численность популяции не может сохраняться слишком большой в течение миллионов поколений. Должно быть, это главная причина того, что наблюдаемая гетерозиготность на локус, кодирующий фермент, для различных организмов редко превышает 30% даже в очень больших популяциях, характерных, например, для насекомых. У млекопитающих, согласно работе [420] (рис. 2), средняя гетерозиготность H у 48 видов распределяется между 0 и 7% почти равномерно, за исключением двух случаев, когда значения H достигают примерно 10%.

Как было показано для рассматриваемого нами гипотетического млекопитающего, в данном нуклеотидном сайте замена основания происходит в среднем через каждые 200 млн. лет. За это время млекопитающее должно пройти через 400 этапов видообразования. Даже продолжительность состояния нейтрального полиморфизма, которая равна $4N_e \pm 2,15$, или $400\,000 \pm \pm 215\,000$ поколений, может иногда превышать продолжительность жизни вида.

Таким образом, процесс стохастического изменения генных частот по обычным понятиям протекает крайне медленно. Предположим, что частота полиморфного аллеля на локус равна p ; тогда ожидаемое (среднее) изменение этой частоты за одно поколение у вида в целом равно нулю, а дисперсия

$$\sigma^2_{dp} = p(1-p)/(2N_e). \quad (8.10)$$

Таким образом, вероятная величина изменения частоты (т. е. σ_{dp}) за одно поколение будет меньше 0,12%. Даже если наблюдения продолжаются более 10 поколений и ведутся за полиморфным аллелем, частота которого равна примерно 50%, вероятная величина изменения частоты составит только около 0,35%.

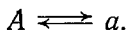
Если вид состоит из нескольких локальных колоний, или де-мо-в, изменения частот в отдельных колониях окажутся, вероятно, более значительными, но в целом будут стремиться к при-

веденной выше величине. Если случайным образом или по неизвестной причине генная частота в отдельной колонии сильно изменится, она рано или поздно приблизится к средней частоте, характерной для вида в целом, вследствие миграции из соседних колоний. Такая ситуация может создать впечатление о наличии сбалансированного полиморфизма. На самом деле полиморфизм в случае нейтральности является временным, но в пределах нашего краткосрочного существования его можно считать практически постоянным.

Отметим, что резкое уменьшение размеров популяции в течение нескольких поколений (например, уменьшение численности некоторых видов дичи в результате слишком интенсивного промысла) вряд ли приводит к значительному уменьшению гетерозиготности (если не считать каких-то экстремальных ситуаций). Для разъяснения этого утверждения предположим, что у нашего гипотетического млекопитающего эффективный размер популяции уменьшался наполовину в каждом поколении в течение последних 10 поколений. Во время этого процесса размер популяции сократился от $N_e = 10^5$ до $N_e = 97,7$, т. е. примерно в тысячу раз, а гетерозиготность должна уменьшиться от $\bar{H}_e = 4,90\%$ до $\bar{H}_e = 4,85\%$, т. е. на незначительную величину.

8.2. Некоторые модели аллельных мутаций

Длительное время в популяционной генетике была принята модель мутирования, согласно которой в локусе существуют два аллеля, а мутации происходят в обоих направлениях, т.е.



Эта модель, несмотря на свою простоту и приемлемость для описания некоторых ситуаций, все же непригодна в общем случае для оценки генетической изменчивости на молекулярном уровне, поскольку общее число возможных аллельных состояний каждого гена крайне велико. Например, для локуса, имеющего 1000 нуклеотидных сайтов, число аллелей, образующихся только в результате замены оснований (исключая структурные изменения), равно 4^{1000} , или примерно 10^{602} . $3 \cdot 1000$ (т. е. 3000 аллелей из этого числа) возникают в результате одноступенчатой замены основания, и, таким образом, вероятность возвращения к исходному аллельному состоянию от любого из них путем одноступенчатой замены составляет всего лишь $1/3000$ (в предположении равной вероятности всех мутационных замен). Суммарное число возможных аллелей почти бесконечно по сравнению с суммарным числом аллелей, имеющих в популяции в любой данный момент времени. Эти соображения были использованы Кимурой и Кроу [246] для обоснования модели, которая позже была названа моделью бесконечного чис-

ла аллелей. Она основана на предположении о таком большом возможном числе изоаллельных состояний локуса, что каждая новая мутация приводит к образованию аллеля, не существовавшего в популяции ранее. Схематично эта модель может быть представлена так:

$$A_1 \longrightarrow A_2 \longrightarrow A_3 \longrightarrow A_4 \longrightarrow A_5 \longrightarrow \dots$$

Если аллели, возникающие в результате мутаций, селективно нейтральны (т. е. селективно равноценны), то уровень гетерозиготности в условиях статистического равновесия, когда образование аллелей путем мутирования уравнивается их случайным вытеснением из популяции, определяется следующим образом.

Пусть v — скорость мутирования на локус на поколение, а f_t — вероятность того, что аллельные состояния двух гомологичных генов особи окажутся идентичными в t -м поколении ($t=0, 1, 2\dots$). Другими словами, f_t — вероятность того, что в t -м поколении особь будет гомозиготной по данному локусу. Предположим, что рассматривается панмиктическая популяция, состоящая из гермафродитных особей; эффективный размер популяции равен N_e , а самооплодотворение происходит с вероятностью $1/N_e$. Тогда

$$f_{t+1} = (1-v)^2 \left[\frac{1}{N_e} \left(\frac{1+f_t}{2} \right) + \left(1 - \frac{1}{N_e} \right) f_t \right]. \quad (8.11)$$

Это выражение можно получить, заметив, что два гомологичных гена будут получены особью $(t+1)$ -го поколения от одной особи или от разных особей t -го поколения с вероятностью $1/N_e$ и $(1-1/N_e)$ соответственно. Если они получены от одной особи, то вероятности их происхождения от одного и того же гена или от двух гомологичных генов равны $1/2$; в этом случае, если они произошли от одного и того же гена, $f=1$, а если от гомологичных, f при переходе к следующему поколению не изменяется. Таким образом, мы получаем первый член в скобках правой части соотношения (8.11). С другой стороны, если эти гены получены от двух разных особей, f также не изменяется. Заметив, что при продолжительном случайном скрещивании гермафродитных особей вероятности идентичности аллелей в обоих случаях равны, получаем второй член выражения, стоящего в скобках. Наконец, наличие множителя $(1-v)^2$ следует из того, что объединение генов в гамете приводит к гомозиготности в следующем поколении только в том случае, когда ни один из этих генов не мутировал. Хотя мы предположили, что популяция состоит из гермафродитных особей, ситуация остается такой же и для «двуполой» популяции, за исключением случаев, когда величина N_e очень мала (см. [67], с. 102).

Предположив, что v очень мало, а N_e очень велико ($v \ll 1 \ll$

$\ll N_e$), уравнение (8.11) можно заменить дифференциальным уравнением

$$\frac{df_t}{dt} = \frac{1}{2N_e} - \left(2v + \frac{1}{2N_e} \right) f_t. \quad (8.12)$$

Отсюда

$$f_t = f_\infty + (f_\infty - f_0) \exp \left[- \left(2v + \frac{1}{2N_e} \right) t \right], \quad (8.13)$$

где

$$f_\infty = \frac{1}{4N_e v + 1} \quad (8.14)$$

— ожидаемая гомозиготность в равновесии.

В реальных популяциях f можно определить как сумму квадратов аллельных частот. Обратная величина была названа Кимурой и Кроу «эффективным числом аллелей» [246]. Средняя гетерозиготность на локус в состоянии равновесия равна

$$\bar{H}_e = 1 - f_\infty = \frac{4N_e v}{4N_e v + 1}. \quad (8.15)$$

Уравнение (8.13) показывает, что если популяция не находится в состоянии равновесия, то она приближается к нему со скоростью $(1 + 4N_e v)/(2N_e)$ за поколение. Это означает, что при очень большой эффективной численности популяции ей требуется огромный промежуток времени (порядка $4N_e$ поколений, если $4N_e v$ невелико) для достижения нового состояния равновесия (например, после резкого увеличения размера популяции во много раз).

Приведенная выше формула для $\bar{H}_e$ определяет среднюю гетерозиготность популяции в состоянии равновесия. Если рассмотреть набор из бесконечного числа эквивалентных популяций, в которых происходят независимые эволюционные изменения вследствие мутирования и генетического дрейфа (с одинаковыми значениями v и N_e), то уровень гетерозиготности в каждой из этих популяций является случайной переменной. Другими словами, величина, равная единице минус сумма квадратов аллельных частот,

$$H_e = 1 - \sum_i x_i^2$$

(где x_i — частота i -го аллеля), может случайно изменяться от популяции к популяции, а $\bar{H}_e$ представляет собой среднее, взятое по всем популяциям в состоянии равновесия. Наблюдаемая гетерозиготность в произвольно выбранной популяции будет скорее всего отличаться от среднего.

Получить дисперсию гетерозиготности значительно сложнее, чем среднее, но тем не менее эта задача была решена Стюартом [532], предложившим формулу

$$\sigma_H^2 = \frac{2\theta}{(1 + \theta)^2 (2 + \theta) (3 + \theta)}, \quad (8.16)$$

где $\theta = 4N_e v$. Вообще говоря, Стюарт [532] вывел соответствующую формулу для модели, предполагающей существование произвольного числа K аллелей, тогда как приведенная выше формула отвечает частному случаю $K = \infty$. Некоторые свойства модели с числом аллелей, равным K , будут обсуждаться ниже.

Интересно также оценить число и относительные частоты различных аллелей в популяции. В рамках модели бесконечного числа аллелей в популяции постоянно происходит их смена, старые аллели со временем заменяются новыми, и в этом смысле стационарного состояния не наблюдается: аллели постоянно находятся в движении. Однако, если сосредоточить внимание на тех аллелях, которые присутствуют в популяции в данный момент времени, и оценить относительное число аллелей, имеющих различные частоты, можно обнаружить определенную закономерность. Это позволяет найти стационарное распределение аллельных частот $\Phi(x)$. Смысл функции $\Phi(x)$ состоит в том, что $\Phi(x)dx$ представляет собой ожидаемое число аллелей, частота которых находится в интервале от x до $x+dx$. Кимура и Кроу [246] показали, что для нейтральных мутаций

$$\Phi(x) = 4N_e v (1-x)^{4N_e v - 1} x^{-1}. \quad (8.17)$$

Отметим, что этот тип распределения частот отличается от вероятностного распределения аллеля данного локуса в эквивалентных популяциях [см. уравнение (8.4) в предыдущем разделе]. Распределение $\Phi(x)$ было названо Юэнсом [97] «спектром частот». В дальнейшем для обозначения величины $4N_e v$ мы будем в основном использовать букву M , хотя иногда для этого применяется буква θ (я постараюсь, чтобы смысл обозначений оставался понятным). Тот факт, что частоты различных аллелей одного локуса в сумме равны единице, можно представить в виде

$$\int_0^1 x \Phi(x) dx = 1.$$

В том, что это действительно так, можно убедиться, подставив вместо $\Phi(x)$ выражение $M(1-x)^{M-1}x^{-1}$ и проинтегрировав левую часть уравнения. Подобным образом средняя величина суммы квадратов аллельных частот, или средняя гомозигот-

ность, равна

$$\bar{f} = \bar{H}_0 = \int_0^1 x^2 \Phi(x) dx = \frac{1}{M+1} \quad (8.18)$$

(где $M = 4N_e v$), что согласуется с (8.14).

Ожидаемое число аллелей, частоты которых находятся в интервале (x_1, x_2) , равно

$$\bar{n}_a(x_1, x_2) = \int_{x_1}^{x_2} \Phi(x) dx. \quad (8.19)$$

Чтобы получить оценку для L эквивалентных локусов, эту величину нужно умножить на L . Одним из важных свойств данного распределения является то, что две независимые переменные, N_e и v , входят в формулу (8.17) в виде произведения $N_e v$, так что распределение $\Phi(x)$ зависит лишь от одного параметра — $M = 4N_e v$. Более того, согласно (8.15), средняя гетерозиготность $\bar{H}_e$ равна $M/(M+1)$, так что, если она известна, величину M можно определить из соотношения

$$M = \bar{H}_e / (1 - \bar{H}_e). \quad (8.20)$$

Таким образом, используя уравнение (8.19), мы можем найти ожидаемое число аллелей в различных частотных классах и сравнить их с соответствующими наблюдаемыми величинами. В качестве примера я приведу данные Уорда и Бердмора [566] по изменчивости белков у камбалы *Pleuronectes platessa*. В этой работе определялась электрофоретически выявляемая изменчивость по 46 локусам, кодирующим белки (39 ферментов и 7 белков, не относящихся к ферментам), в очень больших выборках, взятых из популяции камбалы Бристольского залива. Данная работа, вероятно, является наиболее полным исследованием изменчивости белков у рыб. Полученная в ней средняя гетерозиготность на особь на локус была равна $0,102 \pm 0,026$. Размер выборки варьировал от локуса к локусу: для 8 локусов он превышал 2000, для 9 был равен $2000 \div 500$, для 16 — $500 \div 100$, а для оставшихся 13 локусов был менее 100. Используя уравнение (8.20), получаем $M = 0,114$ для $\bar{H}_e = 0,102$. Ожидаемое число аллелей, частоты которых находятся в интервале (x, x_2) , можно получить, проинтегрировав $L\Phi(x)$, или $46 M(1-x)^{M-1}x^{-1}$, от $x=x_1$ до $x=x_2$. На рис. 8.3 эти величины представлены заштрихованными столбцами, а соответствующие наблюдаемые величины — зачерненными столбцами (что касается столбцов с точками, я объясню их смысл позднее).

Сравнивая наблюдаемое распределение аллельных частот с теоретическим, необходимо помнить, что $\Phi(x)$ относится к аллелям, реально присутствующим в популяции. Возможное

число аллелей бесконечно, однако в данный момент в популяции имеется лишь ограниченное их число, а все остальное огромное количество отсутствующих аллелей нами не рассматривается. Это означает, что каждый мономерный локус вносит в частот-

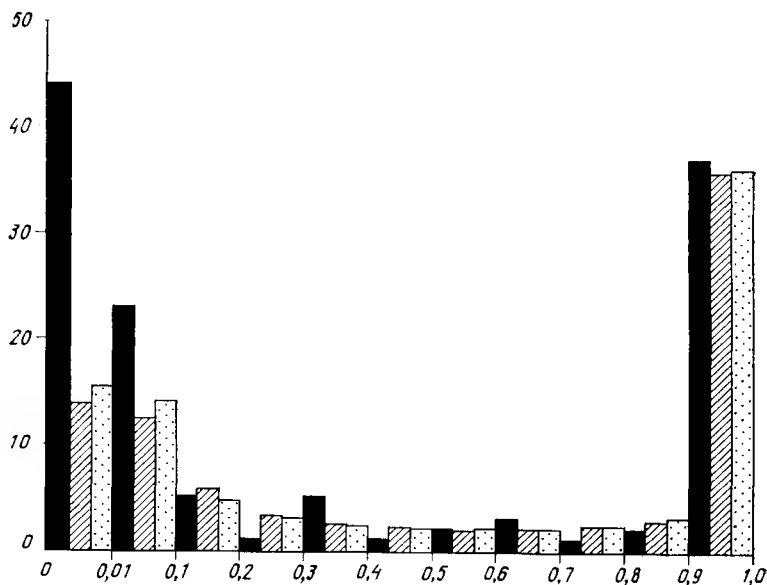


Рис. 8.3. Гистограмма для наблюдаемого и ожидаемых распределений частот аллелей. Черные столбцы — наблюдаемое число аллелей по 46 локусам, кодирующим белки, у камбалы *Pleuronectes platessa* [566]. Заштрихованные столбцы — ожидаемое число аллелей, полученное из модели бесконечного числа аллелей Кимуры и Кроу [246]; столбцы с точками — число аллелей, ожидаемое из модели, в которой предполагается, что скорости мутирования в разных локусах неодинаковы [408]. По оси абсцисс отложена частота аллеля, по оси ординат — число аллелей в различных частотных классах.

ный класс аллелей $x=1,0$ вклад, равный единице. Поэтому я включил мономорфные локусы в крайний правый черный столбец, соответствующий числу аллелей с частотой в интервале $(0,9; 1,0)$. С другой стороны, крайний левый класс, куда относятся аллели с частотой в интервале $(0; 0,01)$, содержит много редких вариантов, найденных лишь в больших выборках. При определении ожидаемого числа аллелей $\bar{n}_a(x < 0,01)$ в этом классе в качестве нижнего предела интегрирования была взята величина, обратная удвоенному среднему размеру выборки на локус:

$$\bar{n}_a(x < 0,01) = \int_{1/(2\bar{n})}^{0,01} \Phi(x) dx = M \int_{1/(2\bar{n})}^{0,01} (1-x)^{M-1} x^{-1} dx, \quad (8.21)$$

где $\bar{n}$ — средний размер выборки, равный для камбалы 678,9 [566]; затем $\bar{n}_a$ умножалась на 46.

Из рис. 8.3 ясно, что ожидаемое число аллелей в данном классе частот гораздо меньше их наблюдаемого числа. Сходная тенденция обнаруживается и для соседнего класса частот (0,01; 0,1). За исключением этих двух классов, соответствие между ожидаемым и наблюдаемым числом аллелей вполне удовлетворительно. Аналогичные данные, иллюстрирующие избыток редких аллелей в популяциях *Drosophila* и человека, описаны также в работах Оты [438, 440]. При анализе данных по популяции камбалы мы считали все 46 изученных локусов эквивалентными в том смысле, что величина $M (=4N_e v)$ для всех них одинакова. Однако намного более вероятным и, конечно, более близким к действительности является предположение, что скорости мутирования нейтральных аллелей неодинаковы в разных локусах из-за различий в размерах генов и в функциональной значимости белков, которые кодируются данными генами. Чтобы учесть эти различия, Ней и др. [408] предложили модель бесконечного числа аллелей, в которую введено предположение о том, что скорость мутирования (точнее, скорость мутирования для нейтральных аллелей), характерная для совокупности локусов, следует гамма-распределению. В этой работе показано, что если для M (или $4N_e v$) характерно гамма-распределение со средней $\bar{M}$ и дисперсией V_M , то функция распределения частот аллелей, задаваемая формулой (8.17), принимает вид

$$\Phi(x) = \frac{\bar{M}x^{-1}(1-x)^{-1}}{[1 - (\bar{M}/\alpha) \ln(1-x)]^{\alpha+1}}, \quad (8.22)$$

где $\alpha = \bar{M}^2/V_M$ — величина, обратная квадрату коэффициента вариации M . Согласно Ней и др., распределения скорости замены аминокислот в ходе эволюции и молекулярных масс соответствующих белков предполагают, что α примерно равно 1 (см. также [122]). Связь между средней гетерозиготностью $\bar{H}_e$ и M в этой модели более сложна, однако Ней [401] вывел полезное соотношение

$$\bar{H}_e = \frac{\bar{M}}{1 + \bar{M}} - \frac{V_M}{(1 - \bar{M})^3}, \quad (8.23)$$

являющееся достаточно хорошей аппроксимацией для $\bar{H}_e$ (за исключением случаев, когда α мало, а $\bar{M}$ — велико). Видно, что неодинаковая скорость мутирования для разных локусов уменьшает среднюю гетерозиготность. Для популяции камбалы $\bar{H}_e = 0,102$, откуда $\bar{M} = 0,128$. Столбцы с точками на рис. 8.3 соответствуют ожидаемому числу аллелей, полученному с по-

мощью формулы (8.22) в предположении, что $\alpha=1$, а $\bar{M}=0,128$. Интересно отметить, что найденные таким образом ожидаемые величины не слишком сильно отличаются от соответствующих величин, полученных из уравнения (8.17) (ср. высоту столбцов с точками и заштрихованных столбцов). Следовательно, избыток редких аллелей имеет место и в том случае, когда учитывается разница в скорости мутирования разных локусов.

Рассмотрим вопрос о редких аллелях, частота которых меньше q , более подробно (применяя теорию к реальным данным, q обычно принимают равным 0,01). Можно показать математически (см. разд. 8.4), что в окрестности точки $x=0$ поведение мутаций, в том числе и мутаций со слабо выраженным благоприятным или неблагоприятным эффектом, вообще говоря, не слишком отличается от поведения селективно нейтральных мутаций. Тогда [403] если $4N_e|s|q$ мало ($|s|$ — абсолютная величина коэффициента отбора мутантного аллеля), то выполняется соотношение

$$\bar{n}_a(x < q) = \int_{1/(2\bar{n})}^q \Phi(x) dx \approx M \ln(2\bar{n}q). \quad (8.24)$$

Здесь M обозначает $4N_e v$, однако v представляет собой скорость мутирования не только селективно нейтральных аллелей, но и всех аллелей, для которых $4N_e|s|q \ll 1$. На самом деле Ней [403] предложил формулу (8.24) для определения скоростей мутирования локусов, кодирующих белки. Далее мы будем обозначать M в этом соотношении через M_q , чтобы отличать его от M в формуле (8.20). Если редкие варианты встречаются более чем в одном локусе, а выборка для каждого локуса велика (средний размер выборки — $\bar{n}$ на локус) и если известно N_e , то суммарная скорость мутирования на локус определяется по формуле

$$v_T = M_q/4N_e, \quad (8.25)$$

где

$$M_q = \bar{n}_a(x < q)/\ln(2\bar{n}q), \quad (8.26)$$

а $q=0,01$.

Поскольку для определения величины M_q по данным о редких вариантах необходимы большие размеры выборки, я отобрал для анализа из работы [566] (см. табл. 2 и 3) 11 локусов, размер выборки для каждого из которых превышает 1000. Средний размер выборки на локус составил $\bar{n}=1956,1$. Из этих 11 локусов 8 являются полиморфными, а 3 — мономорфными. Средняя гетерозиготность для них равна 0,147, т. е. не слишком сильно отличается от соответствующей величины для всех 46 локусов, равной 0,102. Для этих 11 локусов 30 аллелей име-

ли в выборке частоту менее 0,01. Отсюда $\bar{n}_a(x < 0,01) = 30/11 = 2,73$ на локус. Тогда, применив формулу (8.26), где мы предполагаем $q = 0,01$, получаем $M_q = 2,73/\ln(2 \cdot 19,56) = 0,74$. Суммарная скорость мутирования на локус $v_T = 0,74/(4N_e)$. Действительное значение N_e нам неизвестно; если предположить, что оно равно 10^6 , то v_T составит $1,6 \cdot 10^{-7}$.

Здесь важно отметить, что, даже когда мы не знаем реального значения N_e , долю селективно нейтральных (в момент появления) мутаций (обозначим ее $P_{\text{нейтр}}$) можно определить следующим образом. Если теория нейтральности справедлива и если мутантные аллели, обуславливающие полиморфизм белков, селективно нейтральны, то значение $M = 0,114$, которое мы получили путем подстановки наблюдаемой средней гетерозиготности $\bar{H}_e = 0,102$ в формулу (8.20), должно быть равно $4N_e v_0$, где v_0 — скорость мутирования для селективно нейтральных аллелей, выявленных методом электрофореза. Тогда

$$P_{\text{нейтр}} = v_0/v_T = M/M_q. \quad (8.27)$$

Подставляя сюда значения M и M_q , определенные выше, получаем $P_{\text{нейтр}} = 0,15$. Это значит, что в среднем одна мутация из 6,5 является селективно нейтральной, тогда как остальные 5,5 мутаций слишком неблагоприятны, чтобы вносить вклад в полиморфизм белков.

Частота встречаемости редких белковых вариантов была определена также для японского макака (*Macaca fuscata fuscata*) Нозавой и др. (см., например, [424]). Эти авторы проанализировали 32 независимых друг от друга локуса, кодирующих белки, и получили оценку средней гетерозиготности, равную 1,3%, что является довольно низкой величиной даже для млекопитающих. В их обширном исследовании получены следующие результаты (Нозава, личное сообщение, 1981). Среднее число редких вариантов на локус равно $\bar{n}_a(x < 0,01) = 23/22 = 0,719$, а средний размер выборки $\bar{n} = 1609,9$. Отсюда $M_q = \bar{n}_a(x < 0,01)/\ln(2n \cdot 0,01) = 0,207$. Наблюдаемая средняя гетерозиготность на локус $\bar{H}_e = 0,013 \pm 0,0014$, откуда $M = \bar{H}_e/(1 - \bar{H}_e) = 0,0132$. Используя эти величины, получаем $P_{\text{нейтр}} = M/M_q \approx 0,064$. Это означает, что у японских макаков нейтральной является приблизительно одна мутация из 16. Это примерно вдвое меньше соответствующей величины для камбалы.

Численность всей популяции японских макаков равна 20 000—70 000, а по оценкам эффективная численность составляет около одной трети фактической (цитируется по [425]). Если мы примем вслед за Неем $N_e = 2 \cdot 10^4$ [403], то получим $v_T = M_q/(4N_e) = 2,6 \cdot 10^{-6}$. С другой стороны, скорость мутирования нейтральных аллелей равна $v_0 = M/(4N_e) = 1,65 \cdot 10^{-7}$ на поколение.

Эти две оценки $P_{\text{нейтр}}$ (т. е. 0,15 для камбалы и 0,064 для японского макака), — если учесть, что они могут быть определены с большой ошибкой, — предполагают, что 1/10 является вполне разумной оценкой вероятности селективной нейтральности мутации, приводящей к замене аминокислоты. Отметим, что именно эта величина была принята нами в случае гипотетического млекопитающего, рассмотренного в предыдущем разделе.

Оставаясь в рамках модели бесконечного числа нейтральных аллелей, найдем более общее выражение для числа аллелей в популяции. Ожидаемое число различных нейтральных аллелей равно

$$\bar{n}_a = \int_{1/(2N)}^1 \Phi(x) dx = 4N_e v \int_{1/(2N)}^1 (1-x)^{4N_e v - 1} x^{-1} dx. \quad (8.28)$$

Отметим, что нижний предел интегрирования равен не нулю, а $1/2N$ и что $\bar{n}_a$ сильно зависит от N даже в случае, когда $4N_e v$ постоянно: $\bar{n}_a$ увеличивается с увеличением N . Это и понятно, так как, чем больше популяция, тем больше вероятность существования в ней различных редких вариантов. При $N \rightarrow \infty$ $\bar{n}_a$ бесконечно возрастает, если $4N_e v$ остается постоянным. Величину $\bar{n}_a$ можно сопоставить с эффективным числом аллелей $\bar{n}_e$, которое определяется как величина, обратная сумме квадратов частот аллелей. Для нейтральных мутаций ожидаемая величина $\bar{n}_e$ равна

$$\bar{n}_e = 1 \left/ \int_0^1 x^2 \Phi(x) dx \right. = 4N_e v + 1. \quad (8.29)$$

В общем случае эффективное число аллелей (n_e) меньше, чем число разных аллелей, фактически присутствующих в популяции (n_a). Эти две величины совпадают только в том случае, когда частоты различных аллелей одинаковы. Одним из характерных свойств n_e является то, что оно зависит от $4N_e v$, но не зависит от размера выборки n , если только n не очень мало. В качестве примера, иллюстрирующего различие между этими двумя мерами числа аллелей, предположим, что в некой популяции частоты четырех аллелей A_1, A_2, A_3 и A_4 равны 0,5; 0,3; 0,15 и 0,05 соответственно. Тогда сумма квадратов аллельных частот $f = (0,5)^2 + (0,3)^2 + (0,15)^2 + (0,05)^2 = 0,365$, откуда $n_e = 1/f = 2,74$, что меньше, чем действительное число аллелей $n_a = 4$. Ясно, что вклад редких аллелей в величину n_e невелик.

Найдем теперь выражение для числа аллелей в выборке. Предположим, что мы взяли из популяции выборку размера n . Чему будет равно ожидаемое число аллелей в ней? Рассмотрим

аллель, частота которого в популяции равна x . Вероятность того, что этот аллель хотя бы один раз встретится в нашей выборке, есть $1 - (1 - x)^{2n}$. Ожидаемое число аллелей, частоты которых в популяции лежат в интервале от x до $x + dx$, есть $\Phi(x)dx$. Следовательно, ожидаемое число различных аллелей в выборке размера n равно

$$\bar{n}_a = \int_0^1 [1 - (1 - x)^{2n}] \Phi(x) dx. \quad (8.30)$$

Как показал Юэнс [97], в случае селективной нейтральности всех аллелей, когда $\Phi(x)$ задается выражением (8.17), связь между ожидаемым числом аллелей в выборке и популяционным параметром $\theta = 4N_e v$ определяется следующим образом:

$$\bar{n}_a = \frac{\theta}{\theta} + \frac{\theta}{\theta + 1} + \dots + \frac{\theta}{\theta + 2n - 1}. \quad (8.31)$$

Им же показано, что если популяционный параметр θ определять из этой формулы путем подстановки в качестве $\bar{n}_a$ действительного числа различных аллелей, присутствующих в выборке (n'_a), то он будет иметь ряд полезных статистических свойств. Во-первых, такая оценка (обозначим ее $\hat{\theta}$) является оценкой, полученной методом максимального правдоподобия, т. е.

$$n'_a = \frac{\hat{\theta}}{\hat{\theta}} + \frac{\hat{\theta}}{\hat{\theta} + 1} + \dots + \frac{\hat{\theta}}{\hat{\theta} + 2n - 1}. \quad (8.31a)$$

Во-вторых, эта оценка является достаточной, т. е. в n'_a содержится вся информация относительно θ : если дано n'_a , совместное вероятностное распределение числа различных аллелей в выборке не зависит от θ . Юэнс считает, что обычный способ определения θ , основанный на приравнивании $1/(1 + \theta)$ к ожидаемой гомозиготности [см. (8.18)], гораздо хуже, чем данный метод [99]; величина θ , определенная как $1/\bar{f} - 1$, где $\bar{f}$ — сумма квадратов аллельных частот в выборке, сильно смещена и имеет гораздо большую среднеквадратичную ошибку, чем $\hat{\theta}$. Юэнс [97] предложил также приближенные статистические тесты для проверки того, согласуются ли с теорией нейтральности наблюдаемые частоты аллелей в отдельных локусах. Хотя теория Юэнса довольно изящна с точки зрения статистики, мне кажется, что ее применимость к реальным ситуациям весьма ограничена. Как я уже показал, число слабонеблагоприятных аллелей в момент их возникновения, вероятно, в 10 раз превышает число селективно нейтральных аллелей; поэтому, если размер выборки велик (а это необходимо для того, чтобы лю-

бой однолокусный тест, к каким относятся и тесты Юэнса, имел смысл), слабонеблагоприятные аллели увеличивают величину n'_a , что сильно снижает адекватность тестов Юэнса. Этот эффект уже отмечался Неем [401], Отой [442] и позднее Ли [311]. В гл. 9 мы подробнее остановимся на различных статистических тестах, которые предлагались для проверки теории нейтральности с помощью данных по полиморфизму белков.

Все предыдущие рассуждения проводились нами в рамках модели, постулирующей бесконечное число аллельных состояний. Однако иногда удобнее использовать модель, в которой число аллелей равно не бесконечности, а некой конечной величине K . Особенно удобна симметричная модель с K аллелями ($A_1, A_2, A_3, \dots, A_K$), в которых мутации к любому из K аллелей происходят с одинаковой частотой [230].

Пусть v — скорость мутирования на ген на поколение. Обозначим через v_1 скорость мутирования, с которой каждый аллель мутирует к одному из остальных ($K-1$) аллелей, так что $v_1 = v/(K-1)$. Можно показать [230], что при панмиксии ожидаемая гомозиготность в состоянии равновесия (предполагается, что мутации селективно нейтральны) равна

$$\bar{H}_0 = E \left(\sum_{i=0}^K x_i^2 \right) = \frac{1 + (2N_e - 1)(2v_1 - v_1^2 K)}{2N_e - (2N_e - 1)(1 - v_1 K)^2}, \quad (8.32)$$

где x_i — частота i -го аллеля A_i . Пусть $K \rightarrow \infty$, а $(K-1)v_1 = v = \text{const}$; тогда $v_1 \rightarrow 0$, $v_1^2 K \rightarrow 0$, и, следовательно,

$$\bar{H}_0 = \frac{1}{2N_e - (2N_e - 1)(1 - v)^2}. \quad (8.33)$$

Поскольку скорость мутирования $v \ll 1$, знаменатель выражения (8.33) можно принять равным $1 + 4N_e v$, и мы получим формулу, аналогичную (8.14). Кроме того, из (8.32) в предположении $N_e \gg 1 \gg v$ получается полезная формула

$$\bar{H}_0 = \frac{1 + 4N_e v/(K-1)}{1 + 4N_e v K/(K-1)}, \quad (8.33a)$$

согласующаяся с (8.14) при $K \rightarrow \infty$. Среднюю гетерозиготность можно найти по формуле $\bar{H}_e = 1 - \bar{H}_0$.

Фактически величина гомозиготности (сумма квадратов аллельных частот) в данной популяции будет, вероятно, отличаться от среднего значения $\bar{H}_0$ вследствие случайного дрейфа частот. Формула для дисперсии этих отклонений в стационарном состоянии, когда мутационный процесс и случайный дрейф уравнивают друг друга, была получена Стюартом [532]:

$$\text{Var}(H) = \frac{2\theta(1 + \theta/L)}{(1 + \theta + \theta/L)^2 (2 + \theta + \theta/L) (3 + \theta + \theta/L)}, \quad (8.34)$$

где $L=K-1$, а $\theta=4N_e v$. Поскольку $H_e=1-H_0$, то эта формула применима для случая как гомозиготности, так и гетерозиготности. При $L \rightarrow \infty$ (8.34) сводится к формуле (8.16) для дисперсии в модели бесконечного числа аллелей. Стюарт сделал одно интересное замечание, относящееся к нерекуррентному случаю (т. е. $L \rightarrow \infty$): отношение стандартного отклонения σ_H к средней величине $\bar{H}_e$ при $\theta=1$ равно примерно 0,41, однако при уменьшении θ до 0,1 оно увеличивается до 1,75, а при $\theta=0,01$ равно 5,75. Это значит, что при малых $4N_e v$ стандартное отклонение гетерозиготности (H_e) по эквивалентным локусам может быть в несколько раз больше средней величины гетерозиготности.

У гипотетического млекопитающего, рассмотренного нами в предыдущем разделе, средняя гетерозиготность по аллелям, выявляемым методом электрофореза, составляет 4,9%. Отношение $\sigma_H/\bar{H}_e$ в этом случае равно примерно 2,5. Следует отметить, что, даже когда популяционный параметр $\theta=4N_e v$ остается постоянным для данного локуса, фактический уровень гетерозиготности за тот длительный период времени, в течение которого совершается эволюционный процесс, сильно отклоняется от ожидаемой величины ($\bar{H}_e$). Таким образом, излюбленный довод сторонников селекционистской теории, состоящий в том, что уровень гетерозиготности отвечает оптимальной стратегии адаптации вида к условиям среды, в значительной мере теряет свою убедительность.

Возникает естественный вопрос: какой уровень изменчивости будет поддерживаться в популяции конечного размера, если мутации в ней являются гетеротическими, т. е. сверхдоминантными. Кимура и Кроу [246] применили к случаю сверхдоминантных мутаций модель бесконечного числа аллелей. В этой модели предполагается, что все гетерозиготы имеют одинаковую приспособленность, а у гомозигот приспособленность тоже одинакова, но меньше, чем у гетерозигот, на величину s . Другими словами, s является коэффициентом отбора против гомозигот. Ожидается, что в случае сверхдоминирования всех мутантных аллелей в популяции будет поддерживаться значительно большее число аллелей, чем в случае их нейтральности (для исследования этой проблемы Кимура и Кроу разработали приближенный метод). Например, в популяции размером $N_e=10^5$ при скорости мутирования $v=10^{-5}$ и коэффициенте отбора 10^{-3} эффективное число аллелей будет равно 13,5. Тогда средняя гетерозиготность составит $\bar{H}_e=0,926$. Сравним этот результат с тем, который получается в случае селективной нейтральности при такой же скорости мутирования (10^{-5}) и эффективном размере популяции (10^5), когда $n_e=4N_e v+1=5,0$. Средняя гетерозиготность здесь равна $\bar{H}_e=0,80$, т. е. сверхдоминирование

играет сравнительно небольшую роль в ее увеличении, поскольку высокий уровень гетерозиготности уже достигнут за счет мутирования. Если в популяции того же размера $N_e = 10^5$ скорость мутирования ниже, например $\nu = 10^{-6}$, сверхдоминирование становится более эффективным. Например, при $s = 10^{-3}$ эффективное число аллелей в популяции примерно равно 8, тогда как при $s = 0$ оно составляет 1,4.

Хотя предположение о сверхдоминировании многих локусов генома не соответствует действительности, мутантные аллели небольшой части локусов все же могут поддерживаться в популяции благодаря сверхдоминированию. Интересный пример такого рода описан Керром [215], который изучал определяющие пол локусы в популяциях пчел в Бразилии. Пчелы, несущие аллели этих локусов в гаплоидном состоянии, жизнеспособны, однако в гомозиготном состоянии данные аллели летальны, т. е. у самок они проявляют сверхдоминирование с $s = 1$. Для анализа данных по числу определяющих пол аллелей в локальных популяциях пчел Керр применил теорию Кимуры и Кроу [246].

Модель бесконечного числа аллелей со сверхдоминированием рассматривалась также Райтом [600]. Позднее Уоттерсон [572] разработал статистический метод определения отклонений от селективной нейтральности в сторону благоприятного или неблагоприятного эффекта гетерозигот, основанный на оценке гомозиготности (см. также [574, 575]). Как и тест Юэнса [97], тест Уоттерсона основан на выборке по одному локусу. По сравнению с тестами, использующими одновременно много локусов [401, 418, 438, 440], этот метод имеет некоторые недостатки, включая малую статистическую мощность, на что указывал Ли [312].

В молекулярной популяционной генетике существуют и другие заслуживающие внимания модели (например, модель ступенчатого мутирования и модель бесконечного числа сайтов), но мы рассмотрим их в других разделах. В разд. 8.3 и 8.4 я остановлюсь на некоторых важных концепциях стохастической теории популяционной генетики и их математическом анализе.

8.3. Время существования нейтральной мутации в популяции конечного размера

Очень важным параметром, помогающим понять генную изменчивость в популяции, является время существования мутантного аллеля. Проблему его определения мы и рассмотрим в данном разделе. Зададимся вопросом: сколько поколений существует в популяции аллель, ныне имеющий частоту x , с момента своего возникновения в результате мутации? Математическая сторона этой проблемы была разработана Кимурой и

Отой [260], которые решили ее для некоторых простых случаев. В более общем виде она была рассмотрена Маруямой [336] и Маруямой и Кимурой [340].

С тех пор проблема определения времени существования возраста аллеля в популяции продолжала привлекать пристальное внимание специалистов по математической генетике; по этому вопросу был опубликован ряд статей (их обзор см. в [99, 573]). Здесь мы рассмотрим только некоторые простые случаи, имеющие отношение к теме нашей книги.

Если частота нейтрального аллеля после его возникновения была очень низка, а затем увеличилась и в настоящий момент равна x , то среднее время существования этого аллеля, выраженное в числе поколений, определяется как

$$\bar{a}(x) = -4N_e \left(\frac{x}{1-x} \right) \ln x. \quad (8.35)$$

При выводе этой формулы предполагалось, что данный аллель после своего возникновения не мутировал. По терминологии Маруямы [336], таким образом, определяется «время существования аллеля до его фиксации». Предположим, например, что частота нейтрального аллеля равна 10%. Тогда $\bar{a}(0,1) \approx 1,02N_e$, т. е. среднее время существования этого аллеля, выраженное в числе поколений, примерно равно эффективному размеру популяции. Для $x=0,3$ и $0,5$ соответствующие значения $\bar{a}(x)$ равны $\bar{a}(0,3) \approx 2,06N_e$ и $\bar{a}(0,5) \approx 2,77N_e$. Из приведенных примеров следует, что полиморфные аллели (если они селективно нейтральны) поддерживаются в популяции в течение очень долгого времени, которое иногда превышает продолжительность существования вида. При тех же предположениях об отсутствии отбора и повторного мутирования среднее квадратичное время существования (до фиксации) определяется следующим образом:

$$\overline{a^2}(x) = 32N_e^2 \left(\frac{x}{1-x} \ln x - \int_0^x \frac{\ln z}{1-z} dz \right) \quad (8.36)$$

[206]. Для $x=0,1$ эта величина примерно равна $2,8N_e^2$ с дисперсией около $1,8N_e^2$. Это означает, что, хотя время существования (до фиксации) мутантного аллеля, частота которого в данный момент составляет 10%, в среднем равно N_e , в его оценке существует большая неопределенность, причем стандартное отклонение равно приблизительно $1,4N_e$ поколений, что превышает среднее время существования аллеля.

Для того чтобы проиллюстрировать высокую степень асимметричности распределения времени существования, на рис. 8.4 представлены кривые распределения для $x=0,3$ и $0,5$. Эти кри-

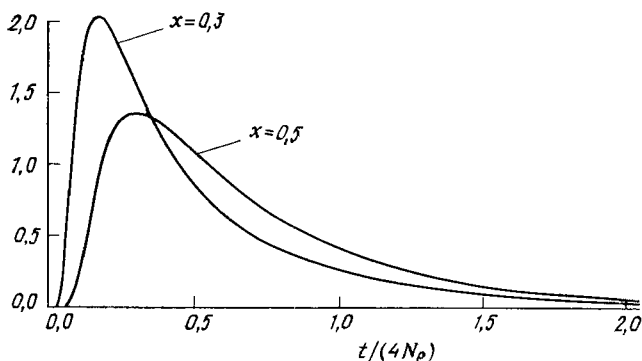


Рис. 8.4. Распределение времени существования аллелей до фиксации. Исходные частоты равны 30 и 50%. По оси абсцисс отложено время, прошедшее до фиксации и измеренное в единицах, равных $4N_e$ поколений; по оси ординат — плотность вероятности. Отметим, что благодаря обратимости рассматриваемого стохастического процесса приведенные кривые представляют собой также распределение времени до элиминации аллелей, частоты которых в данный момент равны 0,3 и 0,5.

вые получены на основании следующих рассуждений. Плотность вероятности времени существования t (возраста) аллеля может быть получена как

$$\lim_{p \rightarrow 0} \left[\phi(p, x; t) \left| \int_0^{\infty} \phi(p, x; t) dt \right. \right],$$

где $\phi(p, x; t)$ задается формулой (3.2). Если время измеряется в единицах, равных $4N_e$, это распределение принимает вид

$$f_x(\lambda) = 2x \sum_{i=1}^{\infty} (2i+1) T_{i-1}^1(z) e^{-t(i+1)\lambda}, \quad (8.37)$$

где $\lambda = t/(4N_e)$, $z = 1 - 2x$, T_{i-1}^1 — полином Гегенбауэра, определяемый как

$$T_{i-1}^1(z) = \frac{i(i+1)}{2} F\left(i+2, 1-i, 2, \frac{1-z}{2}\right), \quad (8.38)$$

где $F(\dots)$ — гипергеометрическая функция. Вычисления можно упростить, приняв во внимание, что для двух первых членов $T_0^1(z) = 0$, $T_1^1(z) = 3z$, а для остальных справедлива следующая рекуррентная формула:

$$iT_i^1(z) = z(2i+1)T_{i-1}^1(z) - (i+1)T_{i-2}^1(z), \quad (8.39)$$

где $i = 2, 3$ и т. д.

При сравнении формулы (8.35) с выражением для среднего времени элиминации мутации обнаруживается интересный

факт: среднее время фиксации аллеля равно среднему времени его элиминации. Действительно, Кимура и Ота [253] показали, что число поколений, прошедших до элиминации нейтрального аллеля, имеющего частоту p (исключая случаи его фиксации), равно

$$\bar{t}_0(p) = -4N_e \left(\frac{p}{1-p} \right) \ln p. \quad (8.40)$$

Это нетрудно понять, поскольку случайный процесс, приводящий к изменению частоты аллеля от p до x , и обратный процесс, благодаря которому частота изменяется от x до p , в отсутствие отбора и мутаций взаимно симметричны.

При применении формулы (8.35) к реальным ситуациям возникает одно досадное затруднение: невозможно установить, увеличилась ли частота аллеля, имеющего в данный момент частоту x , до этой величины от очень низкого значения или, напротив, уменьшилась после состояния фиксации (т. е. частота другого, альтернативного, аллеля увеличилась от очень низкой величины до $1-x$). В последнем случае время существования аллеля в популяции равно $\bar{a}(1-x)$, а не $\bar{a}(x)$. Наиболее разумный выход из положения состоит, как мне кажется, в том, чтобы рассмотреть некую среднюю ситуацию, проводя усреднение путем взвешивания по соответствующим вероятностям. Относительные вероятности данных альтернативных событий равны $1-x$ и x соответственно. Это следует из того, что вероятность достижения нейтральным аллелем, имеющим частоту p , более высокой частоты x до его элиминации или фиксации равна p/x . Аналогично вероятность того, что частота аллеля, имеющего начальную частоту p , увеличится до $1-x$, равна $p/(1-x)$. Объединяя эти два случая, мы получим вероятное время существования аллеля $\bar{A}(x) = (1-x)\bar{a}(x) + x\bar{a}(1-x)$. Отсюда

$$\bar{A}(x) = -4N_e [x \ln x + (1-x) \ln (1-x)]. \quad (8.41)$$

Например, если $x=0,1$, то $\bar{A}(0,1) \approx 1,3N_e$. В этом случае $\bar{A}(0,1)$ лишь немного превышает $\bar{a}(0,1) \approx N_e$. Интересно отметить, что среднее время существования аллеля, вычисленное по формуле (8.41), равно среднему времени, прошедшему до фиксации либо элиминации нейтрального аллеля, имеющего в данный момент частоту x .

До сих пор мы не учитывали возможности повторного мутирования за время изменения генных частот. Если включить в рассмотрение этот процесс, то задача сильно усложнится, поскольку необходимо учитывать не только возможность повторного мутирования до достижения аллелем более высокой частоты, но также и возможность уменьшения частоты уже зафиксированного в популяции мутантного аллеля за счет совме-

стиго влияния на него мутационного процесса и случайного дрейфа. В подобном случае имеет смысл рассматривать «время существования, включающее состояние фиксации» [336]. Пусть v — скорость мутирования на поколение; предположим, что любая возникающая мутация приводит к появлению нового аллеля, ранее в популяции не присутствовавшего. Для селективно нейтрального аллеля можно показать, что если $4N_e v < 1$, то вероятность фиксации этого аллеля

$$u(p) = 1 - (1 - p)^{1-M}, \quad (8.42)$$

где $M = 4N_e v$, а p — начальная частота аллеля. Отметим, что в этом случае фиксация является временной, так как частота аллеля в конечном счете уменьшается из-за мутаций, в результате которых данный аллель превращается в новые аллели. С другой стороны, можно показать, что если $4N_e v \geq 1$, то вероятность достижения аллелем пусть даже временного состояния фиксации равна нулю (предполагается, что N_e велико, так что применима диффузионная модель). В обоих случаях наиболее подходящей мерой является время существования аллеля с учетом времени, в течение которого он находится в состоянии фиксации. Для селективно нейтрального аллеля с частотой x в данный момент времени оно равно

$$\begin{aligned} \bar{a}_{\text{вкл}}(x) = \frac{4N_e}{1-M} \left\{ \int_0^x \frac{1 - (1 - \xi)^{1-M}}{\xi (1 - \xi)^{1-M}} d\xi + \right. \\ \left. + [1 - (1-x)^{1-M}] \int_x^1 \frac{d\xi}{\xi (1 - \xi)^{1-M}} \right\}, \quad (8.43) \end{aligned}$$

где $M = 4N_e v$. В этой формуле предполагается, что начальная частота p очень мала ($p \rightarrow 0$). Формула для $\bar{a}_{\text{вкл}}(x)$ в общем виде приведена в работе Маруямы [336], где представлена также таблица значений этой величины для разных сочетаний $4N_e v$ и x . Время существования аллеля, включающее фиксацию, обычно гораздо больше времени существования до фиксации. Например, при $4N_e v = 0,1$ среднее время существования (включающее фиксацию) аллеля, имеющего в данный момент частоту 10%, составляет около $5,4N_e$ поколений, тогда как время существования этого аллеля до фиксации примерно равно N_e поколениям. Столь большая разница между приведенными величинами объясняется тем, что при выбранном нами темпе мутирования вероятность того, что аллель, частота которого равна 0,1, был когда-то зафиксирован, довольно высока, а для такого аллеля необходимо учитывать как время достижения фиксации, так и продолжительность состояния фиксации и по-

следующего уменьшения частоты. Это и приводит к значительному увеличению среднего времени существования. Отметим, что среднее время существования аллеля, который когда-то находился в состоянии фиксации, равно по меньшей мере $4N_e$ поколениям.

Интересно сравнить время существования аллеля со средним временем достижения мутантным аллелем в первый раз какой-либо частоты (скажем, x), если начальная частота p этого аллеля меньше x . Общая формула для среднего «времени первоначального достижения данной частоты» (first arrival time) была получена Кимурой и Отой [260]. В случае нейтральной мутации в отсутствие повторного мутирования оно равно

$$\bar{t}_x(0) = 4N_e \left[\frac{1-x}{x} \ln(1-x) + 1 \right]. \quad (8.44)$$

При выводе этой формулы предполагалось, что начальная частота аллеля очень низка ($p \rightarrow 0$). Допустим, например, что $x=0,1$. Тогда $\bar{t}_x(0) \approx 0,2N_e$, так что среднее время первоначального достижения данной частоты составляет всего $1/5$ от среднего времени существования аллеля до фиксации. Если $x \ll 1$, (8.44) можно аппроксимировать выражением $\bar{t}_x(0) = 2N_e x$.

Теперь, следуя работе [260], посмотрим, как приведенные выше результаты соотносятся с картиной пространственной генетической изменчивости. Здесь нас особенно интересует следующий вопрос: какой уровень миграции необходим для того, чтобы частоты «редкого» полиморфного аллеля оставались практически одинаковыми в разных локальностях, если этот аллель селективно нейтрален? Прежде всего рассмотрим однородный ареал, образующий окружность радиуса r . Пусть N_T — полное число скрещивающихся особей на поколение; предположим, что они равномерно, с плотностью δ , распределены по ареалу, так что $N_T = 2\pi r \delta$. Обозначим через σ^2 среднеквадратичное расстояние, на которое мигрирует особь за поколение (миграция предполагается изотропной); тогда расстояние миграции за t будет иметь нормальное распределение со средней, равной нулю, и дисперсией $\sigma_t^2 = t\sigma^2$. Расположим множество таких распределений по окружности радиуса $2\pi r$. Если при этом результирующее вероятностное распределение приблизительно равномерно на этой окружности, то частота мутантного аллеля, существующего в популяции t поколений, будет практически одинаковой в разных локальностях. С другой стороны, если с момента появления в популяции мутантного аллеля прошло мало времени, а результирующее вероятностное распределение заметно отличается от равномерного, то в популяции возникает четко выраженная локальная дифференциация аллельных частот.

Заменяв t на $\bar{a}(x)$ и учитывая, что полученное результирующее распределение практически равномерно при $2\pi r \leq 2\sigma_t$, или

$$\pi r \leq \sigma_t, \quad (8.45)$$

имеем

$$N_T \leq 2\delta\sigma \sqrt{\bar{a}(x)}. \quad (8.46)$$

что является условием равномерного пространственного распределения частот нейтральных аллелей.

Рассмотрим теперь двумерный ареал, образующий сферу радиуса r . Предположим, что особи распределены по ареалу равномерно с плотностью δ . Пусть σ^2 — среднеквадратичное расстояние миграции особей в одном поколении, причем миграция изотропна. Рассматривая на сфере двумерное нормальное распределение расстояния миграции в t поколениях и полагая, что полученное вероятностное распределение на сфере сходно с рассмотренным выше распределением, с учетом (8.45) получаем

$$\pi N_T \leq 4\delta\sigma^2 t \quad (8.47)$$

для двумерного ареала, так как в этом случае $N_T = 4\pi r^2 \delta$.

Для мутантного аллеля, средняя частота которого во всей популяции равна 0,1, условие равномерного распределения приближенно сводится к

$$\sigma^2 \delta \geq 1, \quad (8.48)$$

если подставить в (8.47) $t = \bar{a}(0,1) \approx N_T$.

Если популяция подразделена на колонии (субпопуляции), в каждой из которых имеет место случайное скрещивание, а миграция в пределах одного поколения ограничена смежными колониями (двумерная модель «степпинг-стоун» (stepping stone) Кимуры и Вейсса [266]), условие (8.48) превращается в

$$mN \geq 1, \quad (8.49)$$

где m — скорость, с которой каждая колония в каждом поколении обменивается особями с четырьмя прилежащими к ней колониями, а N — эффективный размер колонии. Однако, поскольку время существования мутации $\bar{a}(x)$ при $x=0,1$ имеет довольно большое стандартное отклонение, для получения необходимого условия лучше использовать в качестве t в уравнении (8.47) время первоначального достижения данной частоты $\bar{t}_x(0)$ при $x=0,1$, что приближенно дает

$$mN \geq 5. \quad (8.50)$$

Эти результаты в основном совпадают с выкладками Кимуры и Маруямы [251], основанными на совсем других предпосыл-

ках. Для нейтрального мутантного аллеля, имеющего частоту $x=0,05$, соответствующая формула принимает вид

$$mN \geq 10. \quad (8.51)$$

Это означает, что равномерное распределение генных частот в колониях достигается в том случае, когда между соседними колониями происходит обмен особями, составляющий в среднем по крайней мере 2,5 особей на поколение независимо от размера колонии. Таким образом, мы приходим к следующему выводу: когда средняя частота «редкого» полиморфного аллеля составляет во всей популяции несколько процентов, частоты этого аллеля в различных локальностях будут по существу одинаковыми, если данный аллель селективно нейтрален и если между смежными колониями происходит миграция со средней скоростью, равной по крайней мере нескольким особям на поколение. Вероятно, у многих видов *Drosophila*, популяции которых имеют огромный размер, а особи отличаются высокой подвижностью и, кроме того, способны мигрировать на далекие расстояния с потоками воздуха, обычно достигается фактическая панмиксия даже в том случае, когда отдельные популяции расположены очень далеко друг от друга.

Кроме того, равномерность пространственного распределения частот нейтральных генов по крайней мере частично поддерживается благодаря ассоциативному сверхдоминированию в нейтральном локусе, обусловленному сцеплением с ним селективных локусов [433, 452, 456]. Это приводит к своего рода инерции, создаваемой ассоциативным сверхдоминированием, вследствие чего при временном изменении локальной частоты аллеля последняя стремится вернуться к прежнему значению, несмотря на отсутствие в случае нейтральности определенной равновесной частоты, поддерживающейся в течение долгого времени [436].

Исследуя структуру генетической изменчивости в субпопуляциях *Drosophila pseudoobscura*, Пракаш, Левонтин и Хабби [473] отвергли модель нейтральных изоаллелей на том основании, что частоты редких аллелей нескольких локусов (таких, как локус, кодирующий малатдегидрогеназу) в популяциях, разделенных большими расстояниями, были практически одинаковыми; по мнению этих авторов, гипотеза о нейтральности изоаллелей не в состоянии объяснить подобную идентичность частот. Приведенные выше рассуждения, как мне кажется, показывают, что результаты работы [473] не противоречат теории нейтральности.

Мне хочется также отметить, что в случае, когда наблюдаемое равномерное распределение частот редких аллелей в локальных субпопуляциях обусловлено не миграцией, а слабым «балансирующим» отбором, эффективный размер локальной

популяции (но не вида в целом) должен, по всей вероятности, быть порядка миллиона, не говоря уже о том, что в этом случае коэффициенты отбора в разных локальностях должны быть равными. Это происходит потому, что, как было впервые показано Робертсоном [487] для сверхдоминантных аллелей, балансирующий отбор не замедляет, а ускоряет фиксацию, если равновесные частоты не лежат в интервале 0,2—0,8, а $N_e(s_1+s_2)$ не очень велико (s_1 и s_2 — коэффициенты отбора против двух гомозигот). Действительно, при равновесной частоте 5% величина $N_e(s_1+s_2)$ должна составлять 2100, чтобы произошло замедление фиксации в 100 раз по сравнению со случаем полной нейтральности (см. также [67], с. 414).

В приведенном выше обсуждении пространственного распределения нейтральных аллелей, основанном на представлении о времени существования мутации, мы неявно предполагаем, что между колониями осуществляется достаточно интенсивная миграция, чтобы система была по существу эквивалентна панмиктической популяции. При очень низком уровне миграции проблема становится более сложной. Маруяма [331, 333] показал, что некоторые параметры, например вероятность фиксации мутантного гена и суммарная гетерозиготность по определенной мутации в процессе ее фиксации или элиминации, географически инвариантны. Другими словами, эти параметры не зависят от географического распределения при условии, что ни одна часть популяции не изолирована полностью от остальных (см. также [337]). Весьма возможно, что по сравнению с этими величинами время существования мутации обнаруживает большую зависимость от степени подразделенности популяции, в связи с чем требуются дальнейшие исследования.

Время существования аллелей с выраженным неблагоприятным эффектом гораздо меньше, чем у нейтральных аллелей, поэтому в данном случае наличие локальной дифференциации генных частот более вероятно.

8.4. Интеграл от произвольной функции вдоль траектории изменения частоты гена

В предыдущем разделе мы определили время существования в конечной популяции аллеля, имеющего данную частоту. С математической точки зрения этот вопрос является частным случаем гораздо более общей задачи определения суммы (т. е. интеграла) произвольной функции генной частоты x , скажем $f(x)$, вдоль траектории изменения частоты гена в выборке. В гл. 3 мы привели несколько примеров таких траекторий (см. рис. 3.3), полученных путем моделирования методом Монте-Карло поведения нейтрального аллеля в популяции малого размера.

При математическом анализе стохастического поведения мутантного аллеля в конечной популяции часто бывает удобно рассматривать набор траекторий изменения генных частот в выборках. В такой постановке задача определения среднего времени фиксации мутантного аллеля сводится к вычислению среднего значения интеграла вдоль тех траекторий, которые, начинаясь с исходной частоты $x=p$, приводят к фиксации ($x=$

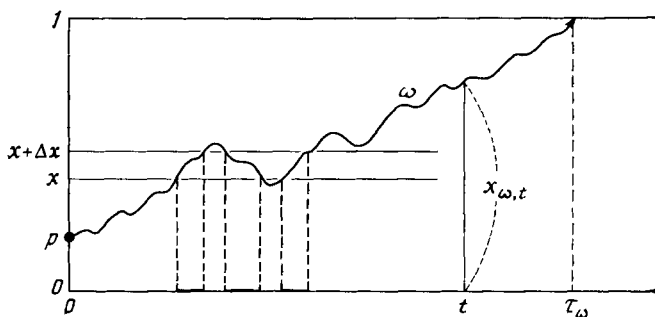


Рис. 8.5. Траектория выборки ω , иллюстрирующая изменение частоты гена во времени. Она начинается в момент времени 0 с частоты p и приводит к фиксации в момент времени τ_ω . Сплошными горизонтальными отрезками указаны промежутки времени, когда частота аллеля находится в интервале частот $(x; x+\Delta x)$. По оси ординат отложена частота гена, по оси абсцисс — время.

$=1$), причем в качестве подынтегрального выражения берется $f(x)=1$. Сходным образом ожидаемая величина суммарной доли гетерозигот, порождаемых мутантным аллелем в панмиктической популяции до его фиксации или элиминации, равна среднему значению интеграла $f(x)=2x(1-x)$ вдоль всех траекторий, начинающихся с частоты p и заканчивающихся фиксацией или потерей мутации.

Хотя детальный анализ этого вопроса выходит за рамки данной книги, мне хотелось бы вкратце рассмотреть его чисто теоретически, уделив особое внимание «времени пребывания» («sojourn time») аллеля в различных классах частот в ходе изменения его частоты.

Рассмотрим сначала набор траекторий выборок, которые начинаются с исходной частоты p и приводят к фиксации аллеля или к его элиминации. Обозначим через $x_{\omega,t}$ частоту этого аллеля в момент времени t для некой траектории ω (см. рис. 8.5). Пусть $F(p)$ — ожидаемая величина суммы $f(x)$ вдоль этой траектории, т. е.

$$F(p) = E \left[\int_0^{\tau_\omega} f(x_{\omega,t}) dt \right], \quad (8.52)$$

где E — оператор математического ожидания (среднего), определяемого по всем траекториям выборок, а τ_ω — время, за которое данная траектория приводит к фиксации или потере. $F(p)$ можно выразить через плотность вероятности перехода $\phi(p, x; t)$ следующим образом:

$$F(p) = \int_0^\infty dt \int_0^1 \phi(p, x; t) f(x) dx, \quad (8.53)$$

где $\phi(p, x; t)$ — плотность вероятности того, что частота данного аллеля станет равной x в момент времени t , если при $t=0$ она была равна p . Можно показать, что $F(p)$ удовлетворяет следующему дифференциальному уравнению:

$$\frac{1}{2} V_{\delta p} \frac{d^2 F(p)}{dp^2} + M_{\delta p} \frac{dF(p)}{dp} + f(p) = 0, \quad (8.54)$$

где $M_{\delta p}$ и $V_{\delta p}$ — соответственно среднее и дисперсия изменения генной частоты p за одно поколение.

Если пренебречь появлением мутаций в этом процессе изменения частоты, то две границы, $x=0$ и $x=1$, будут являться поглощающими состояниями, а соответствующие граничные условия будут

$$F(0) = F(1) = 0. \quad (8.55)$$

Решение уравнения (8.54) при этих граничных условиях имеет вид [338]

$$F(p) = [1 - u(p)] \int_0^p \psi_f(\xi) u(\xi) d\xi + u(p) \int_p^1 \psi_f(\xi) [1 - u(\xi)] d\xi, \quad (8.56)$$

где

$$\psi_f(\xi) = 2f(\xi) \int_0^1 G(\lambda) d\lambda / [V_{\delta\xi} G(\xi)], \quad (8.56a)$$

$$u(p) = \int_0^p G(\xi) d\xi / \int_0^1 G(\xi) d\xi, \quad (8.56b)$$

$$G(\xi) = \exp \left[-2 \int_0^\xi (M_{\delta x} / V_{\delta x}) dx \right]. \quad (8.56в)$$

В этих формулах $u(p)$ — вероятность фиксации мутантного аллеля, исходная частота которого была равна p [227]; $\exp(\cdot)$ — экспоненциальная функция. Отметим, что в частном случае, когда мутантный аллель обладает селективным преимуществом s , $M_{\delta p} = sp(1-p)$, откуда в сочетании с $V_{\delta p} = p(1-p)/(2N_e)$

получаем из (8.56 в) $G(\xi) = e^{-S\xi}$, где $S = 4N_e s$. Подставляя это в (8.56б), мы приходим к выражению $u(p) = (1 - e^{-Sp}) / (1 - e^{-S})$, что согласуется с уравнением (3.11).

Сумма долей гетерозигот по всем поколениям до фиксации или элиминации мутантного аллеля, начальная частота которого равна p , определяется из уравнения (8.56) путем подстановки $f(p) = 2p(1-p)$, что соответствует ожидаемой частоте гетерозигот по мутантному аллелю, частота которого в панмиктической популяции равна p .

Предположим, что данный мутантный аллель имеет селективное преимущество s и что в процессе изменения частот мутаций не происходит. В этом случае $G(\xi) = e^{-S\xi}$, $V_{\delta\xi} = \xi(1-\xi) / (2N_e)$, $f(\xi) = 2\xi(1-\xi)$, $\psi_f(\xi) = 8N_e s \xi(1-e^{-S}) / S$. Обозначив ожидаемую величину суммарной доли гетерозигот через $H(p)$, мы получим $H(p) = 2[u(p) - p] / s$, что согласуется с уравнением (3.15).

Для селективно нейтрального мутантного аллеля при $s \rightarrow 0$ уравнение для $H(p)$ сводится к $H(p) = 4N_e p(1-p)$. Тогда полная гетерозиготность по одной мутации равна

$$H(1/2N) \approx 2N_e / N. \quad (8.57)$$

При $N_e = N$, т. е. когда эффективный размер популяции равен ее фактическому размеру, эта величина равна 2.

Среднее время, прошедшее до фиксации или утраты мутации, обозначенное нами $T(p)$, можно получить путем подстановки $f(p) = 1$ в уравнение (8.56). Для селективно нейтральной мутации $M_{\delta x} = 0$, откуда $G(\xi) = 1$, $u(p) = p$ и, следовательно, $\psi_f(\xi) = 4N_e / [\xi(1-\xi)]$. Таким образом,

$$T(p) = -4N_e [(1-p) \ln(1-p) + p \ln p], \quad (8.58)$$

что согласуется с результатами, полученными Уоттерсоном [570].

Обратимся теперь к задаче определения средней величины суммы $f(x)$ вдоль траекторий изменения генных частот в выборках при условии их фиксации, т. е. будем определять среднюю только для траекторий, заканчивающихся фиксацией, не рассматривая траекторий, ведущих к потере аллеля. Говоря на языке математики, нам нужно рассмотреть величину

$$E \left[\int_0^{\tau_\omega} f(x_\omega, t) dt \mid x_\omega, \tau_\omega = 1 \right].$$

Через плотность вероятности перехода эта величина выражается как $F_1(p)/u(p)$, где $u(p)$ — вероятность фиксации, а

$$F_1(p) = \int_0^\infty dt \int_0^1 \phi(p, x; t) f(x) u(x) dx. \quad (8.59)$$

В этом выражении индекс 1 обозначает, что речь идет о траекториях, ведущих к фиксации ($x=1$). Можно показать, что $F_1(p)$ удовлетворяет дифференциальному уравнению

$$\frac{1}{2} V_{\delta p} F_1''(p) + M_{\delta p} F_1'(p) + u(p) f(p) = 0 \quad (8.60)$$

($'$ и $''$ означают дифференцирование). Это уравнение аналогично (8.54). Единственное различие между ними заключается в том, что в уравнение (8.60) вместо $f(p)$ входит $u(p)f(p)$.

В отсутствие мутаций граничными условиями являются

$$F_1(1) = F_1(0) = 0. \quad (8.61)$$

Тогда ожидаемая величина суммы $f(x)$ вдоль траектории выборки, начинающейся с p и кончающейся фиксацией, равна

$$\begin{aligned} E \left[\int_0^{\tau_\omega} f(x_{\omega,t}) dt \mid x_{\omega,\tau_\omega} = 1 \right] = \\ = \frac{1-u(p)}{u(p)} \int_0^p \psi_f(\xi) [u(\xi)]^2 d\xi + \int_p^1 \psi_f(\xi) u(\xi) [1-u(\xi)] d\xi, \end{aligned} \quad (8.62)$$

где $\psi_f(\xi)$ задается формулой (8.56а).

Среднее время фиксации (исключая случаи возможной элиминации аллеля) селективно нейтральной мутации можно получить из этого выражения, приняв $f(x)=1$ и заметив, что поскольку для нейтральных аллелей $G(\xi)=1$, то $\psi_f(\xi) = 4N_e / [\xi(1-\xi)]$. Подставив эти выражения в (8.62), мы сразу же получим, что в предельном случае $p \rightarrow 0$ оно сводится к $4N_e$ в соответствии с [253].

Приведенный выше анализ относится к среднему значению (моменту первого порядка) стохастического интеграла. Теория для нахождения моментов более высоких порядков разработана Маруемой и Кимурой; читателей, которых интересует этот вид задач, мы отсылаем к их работе [338], а также к работам [337, 340, 395].

Перейдем теперь к обсуждению вопросов, относящихся к так называемому «времени пребывания» (sojourn time), т. е. продолжительности пребывания мутантного аллеля в различных классах частот в процессе изменения его частоты. Особое внимание мы уделим среднему времени пребывания для траекторий, заканчивающихся фиксацией (см. рис. 8.5). Используя общую теорию, изложенную в настоящем разделе, среднее время пребывания аллеля в состоянии с данной частотой (равной, скажем, y) можно найти, приняв $f(x) = \delta(x-y)$, где $\delta(\cdot)$ — дельта-функция Дирака. Эта функция определяет вероятность, сосредоточен-

ную в одной точке $x=y$, а для всех $x \neq y$ она равна нулю. В более общем виде для любой функции $\phi(x)$ мы имеем

$$\int \phi(x) \delta(x-y) dx = \phi(y), \quad (8.63)$$

при условии, что область интегрирования включает точку y . В дальнейшем, однако, генную частоту мы будем обозначать не через y , а через x .

Пусть $\Phi_1(p, x)$ — плотность вероятности среднего времени пребывания для траекторий, начинающихся с частоты p и приводящих к фиксации. Это означает, что $\Phi_1(p, x) dx$ представляет собой среднее значение полного времени, в течение которого частота аллеля находится в интервале $(x; x+dx)$ до фиксации аллеля в популяции (см. жирные отрезки на оси абсцисс рис. 8.5). Из сказанного выше очевидно, что $\Phi_1(p, x)$ задается уравнением (8.62), если в (8.56а) положить $f(\xi) = \delta(\xi - x)$. Таким образом, мы получаем для $\Phi_1(p, x)$ следующие выражения:

$$\left. \begin{aligned} \Phi_1(p, x) &= \frac{2u(x)[1-u(x)]}{V_{\delta x} u'(x)} \quad (x \geq p) \\ \Phi_1(p, x) &= \frac{2[1-u(p)]u^2 x}{u(p)V_{\delta x} u'(x)} \quad (x < p) \end{aligned} \right\}, \quad (8.64)$$

где $u(x)$ — вероятность фиксации, а

$$u'(x) = du(x)/dx = G(x) \left| \int_0^1 G(x) dx \right|.$$

Для впервые возникающего мутантного аллеля $p=1/(2N)$, где N — фактическая численность популяции. Из двух выражений (8.64) нам подходит только то, для которого $x \geq p$, а среднее число поколений, в течение которых частота мутантного аллеля до фиксации находится в интервале от x до $x+dx$, равно $\Phi_1(x) dx$, где

$$\Phi_1(x) = \frac{2u(x)[1-u(x)]}{V_{\delta x} u'(x)} \quad \left(1 > x \geq \frac{1}{2N}\right). \quad (8.64a)$$

Среднее время пребывания в каждом классе частот $x=i/(2N)$ ($i=1, 2, \dots, 1-1/2N$) определяется как $\Phi_1(x)/(2N)$.

Если мутация селективно нейтральна, то $u(x)=x$, $V_{\delta x} = x(1-x)/(2N_e)$, и приведенная выше формула сводится к

$$\Phi_1(x) = 4N_e. \quad (8.65)$$

Следовательно, нейтральный аллель до фиксации проводит в каждом классе частот в среднем $\Phi_1(x)/2N = 2N_e/N$ поколений. Суммируя по всем $(2N-1)$ классам частот, получаем, что среднее время фиксации нейтральной мутации примерно равно $4N_e$

поколениям [см. (3.23)]. Если фактическая численность популяции равна ее эффективной численности, то время пребывания мутантного аллеля в каждом классе частот на его пути к фиксации составляет в среднем 2 поколения.

Рассмотрим теперь случай, когда мутации испытывают давление отбора. Пусть s — селективное преимущество мутантного аллеля по сравнению с ранее существовавшим аллелем. Тогда $M_{\delta x} = sx(1-x)$, $V_{\delta x} = x(1-x)/(2N_e)$ и $u(x) = (1-e^{-Sx})/(1-e^{-S})$; отсюда

$$\Phi_1(x) = \frac{(1-e^{-Sx})[1-e^{-S(1-x)}]}{sx(1-x)(1-e^{-S})}, \quad (8.66)$$

где $S=4N_es$. В окрестности $x=0$ правая часть уравнения (8.66) становится примерно равной $4N_e$. Это означает, что даже аллель, обладающий явным селективным преимуществом или напротив невыгодностью (величина $|s|$ предполагается малой, однако $4N_e|s|$ может быть большим), ведет себя как селективно нейтральный, пока он еще достаточно редок в популяции. На этой стадии судьба такого аллеля определяется в основном случайным генетическим дрейфом.

При промежуточных частотах среднее время пребывания для явно благоприятной мутации ($4N_es \gg 1$) меньше, чем для нейтрального аллеля. Для x , при которых $4N_esx$ и $4N_es(1-x)$ достаточно больше 1 (скажем, больше 5), справедлива приближенная формула

$$\Phi_1(x) dx = \frac{dx}{sx(1-x)}. \quad (8.67)$$

Из детерминистического анализа изменения генных частот хорошо известно [см. уравнение (6.4)], что скорость изменения частоты аллеля равна

$$\frac{dx}{dt} = sx(1-x), \quad (8.68)$$

так что время, за которое частота мутантного аллеля, имеющего селективное преимущество s (>0), увеличится от x до $x+dx$, равно $dt=dx/sx(1-x)$. Это означает, что в данном интервале частот поведение мутантных аллелей можно удовлетворительно описать с помощью детерминистской теории. В окрестности $x=1$ правая часть (8.66) снова становится примерно равной $4N_e$; следовательно, селективно благоприятный мутантный аллель при приближении к точке фиксации вновь начинает вести себя как нейтральный. Тогда детерминистическое уравнение (8.68) опять становится неприменимым, и нужно использовать уравнение (8.66). Коэффициент отбора s может быть как положительным, так и отрицательным; в последнем случае удобно сделать

замену $s = -s'$, где $s' > 0$. После подстановки в (8.66) получим

$$\Phi_1(x) = \frac{(1 - e^{-S'x}) [1 - e^{-S'(1-x)}]}{s'x(1-x)(1 - e^{-S'})}, \quad (8.69)$$

где $S' = 4N_e s'$. Сравнение соотношений (8.66) и (8.69) показывает, что два мутантных аллеля, один из которых имеет коэффициент отбора s , а другой $-s$, характеризуются одинаковым временем пребывания в различных классах частот при условии, что в конце концов эти аллели фиксируются. Более того, сравнение упомянутых выше формул с (8.65) показывает, что среднее время пребывания для этих аллелей меньше, чем для нейтральной мутации. Таким образом, среднее время фиксации $\int_p^1 \Phi_1(x) dx$ [где $p = 1/(2N)$] становится меньше как для вредных, так и для благоприятных мутаций. Совершенно очевидно, что время фиксации благоприятного аллеля меньше, чем нейтрального, поскольку увеличению частоты первого способствует естественный отбор, однако довольно трудно представить себе, что оно будет меньше и для селективно неблагоприятной мутации, увеличение частоты которой сдерживается отрицательным отбором. Этот кажущийся парадокс можно разрешить, если учесть, что вероятности этих двух событий совершенно различны, а именно: фиксация неблагоприятной мутации — событие очень редкое; такая мутация может достичь фиксации только в результате счастливого стечения обстоятельств, способствующих случайному увеличению ее частоты. Другими словами, мутация случайно фиксируется в популяции вопреки направленному против нее отбору лишь в случае серии удачных «прыжков»; такой процесс должен быстро закончиться.

До сих пор мы рассматривали среднее время пребывания для траекторий, приводящих к фиксации (аналогичный метод используется в случае траекторий, приводящих к потере аллеля). Помимо этого весьма полезно получить выражения для среднего времени пребывания по всем траекториям, вместе взятым (т. е. по траекториям, заканчивающимся как фиксацией, так и элиминацией аллеля, имеющего начальную частоту p). Плотность вероятности $\Phi(p, x)$ этого времени пребывания удовлетворяет уравнению (8.54) при $f(p) = \delta(x - p)$. Предположим, что как $x = 0$, так и $x = 1$ — это поглощающие состояния; тогда граничные условия имеют вид $\Phi(0, x) = \Phi(1, x) = 0$, а решение выражается следующим образом:

$$\left. \begin{aligned} \Phi(p, x) &= \frac{2u(p) [1 - u(x)]}{V_{\delta x} u'(x)} & (x \geq p) \\ \Phi(p, x) &= \frac{2[1 - u(p)] u(x)}{V_{\delta x} u'(x)} & (x < p) \end{aligned} \right\}. \quad (8.70)$$

Для селективно нейтрального аллеля эти соотношения сводятся к

$$\left. \begin{aligned} \Phi(p, x) &= 4N_e p/x & (x \geq p) \\ \Phi(p, x) &= 4N_e (1-p)/(1-x) & (x < p) \end{aligned} \right\}. \quad (8.71)$$

Соотношения (8.70) эквивалентны «переходной функции» («transient function») $t(x)$ Юэнса [96], а также функции распределения стационарного потока (steady flux distribution) при $v_m=1$ [232]. Для мутантного аллеля, впервые появляющегося в популяции, плотность вероятности среднего времени пребывания его на частоте x до фиксации или элиминации можно получить из первого соотношения (8.70). В этом случае, приняв во внимание, что

$$u\left(\frac{1}{2N}\right) \approx \frac{1}{2N} \cdot \frac{1}{\int_0^1 G(x) dx}, \quad (8.72)$$

получаем

$$\Phi(x) = \frac{1}{NV_{\delta x} G(x)} \cdot \frac{\int_x^1 G(\xi) d\xi}{\int_0^1 G(\xi) d\xi} \quad \left(\frac{1}{2N} \leq x < 1\right), \quad (8.73)$$

где $G(\xi)$ задается формулой (8.56в). В окрестности $x=0$ функция $\Phi(x)$ примерно равна $1/(NV_{\delta x})$. Тогда, если положить $V_{\delta x} \approx x/(2N_e)$,

$$\Phi(x) \approx (2N_e/N) \cdot x^{-1} \quad (0 < x \ll 1). \quad (8.74)$$

Сравнение этого выражения с формулой для среднего времени пребывания в случае нейтральной мутации [см. (8.71)], где $p=1/(2N)$, показывает, что в окрестности $x=0$ поведение мутаций в общем случае, включая слабоблагоприятные или слабонеблагоприятные, по существу совпадает с поведением нейтральной мутации.

С помощью общего метода, описанного в данной главе, можно детально исследовать поведение сверхдоминантных аллелей. Читателей, интересующихся этим вопросом, мы отсылаем к работе [417].

Мы ограничились здесь рассмотрением среднего времени пребывания, т. е. момента первого порядка, однако, как показано Маруемой и Кимурой [338] и Нагилаки [395], можно получить и моменты более высоких порядков. Нагилаки нашел не только общее выражение для момента n -го порядка, но и вероятностное распределение времени пребывания (более подробно см. в [395]).

8.5. Модель ступенчатого мутирования

Рассмотренная нами в разд. 8.2 модель с бесконечным числом аллелей основана на предположении, что каждая новая мутация порождает новый аллель, не существовавший в популяции ранее. Эта модель была бы вполне адекватной, если бы отдельные варианты в популяции можно было идентифицировать на уровне нуклеотидных или аминокислотных сайтов. Однако в настоящее время экспериментальный анализ генетической изменчивости по-прежнему основывается на гораздо более грубых методах идентификации электрофоретически выявленных белковых

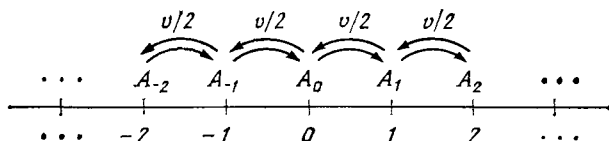


Рис. 8.6. Модель ступенчатого мутирования.

вариантов. Мутация выявляется только в том случае, когда сопровождается заменой аминокислоты, приводящей к изменению электрофоретической подвижности белка. Аллели идентифицируются только как дискретный набор полос в гелях.

Учитывая такую низкую разрешающую способность метода выявления мутаций, Кимура и Ота [259] предложили модель ступенчатого возникновения нейтральных мутаций; некоторые математические следствия этой модели были впервые получены Отой и Кимурой [457]. В своей наиболее простой форме эта модель предполагает, что вся совокупность аллелей данного локуса может быть представлена в виде некой последовательности. Если аллель изменяется вследствие мутации (т. е. вследствие замещения аминокислоты, приводящего к изменению электрического заряда белковой молекулы), то его местоположение в последовательности перемещается на один шаг в положительном или отрицательном направлении (см. рис. 8.6). В рамках этой модели одно положительное и одно отрицательное изменение компенсируют друг друга, так что аллель возвращается к своему первоначальному состоянию.

Ранее Балмер [33] подверг теорию нейтральности критике на том основании, что существует определенная связь между частотами аллелей и их относительными электрофоретическими подвижностями. Анализируя данные Пракаша и др. [473], он отметил следующий интересный факт: подвижность наиболее часто встречающихся в популяции аллелей занимает промежуточное положение между подвижностями редких аллелей. Другими словами, если обозначить каждый часто встречающийся аллель буквой С, а каждый редкий — буквой R, схема располо-

жения этих аллелей в геле (ее иногда называют электрофоретическим профилем) будет всегда RCR, а не CRR или RCC. На основании этих результатов Балмер пришел к выводу, что данные аллели не являются селективно нейтральными, поскольку в случае нейтральности все возможные случаи их расположения должны быть равновероятными.

Смит [349], однако, указал на то, что результаты Пракаша не противоречат теории нейтральности и поэтому критика Балмера несостоятельна. Кроме того, Кимура и Ота [253], используя стохастическую теорию популяционной генетики и основы-

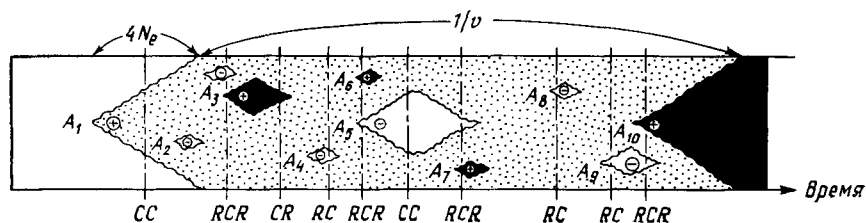


Рис. 8.7. Типичная картина утраты и распространения селективно нейтральных электрофоретически различных мутаций в популяции конечного размера (по [259], с изменениями). Предполагается, что мутации возникают со скоростью одна мутация на каждые 10 поколений ($4N_0v=0,2$). Подробности см. в тексте.

ваясь на природе мутаций, выявляемых методом электрофореза, объяснили, почему отмеченное Балмером явление представляет собой естественное следствие нейтральности полиморфизма белков.

На рис. 8.7 представлена типичная картина утраты и распространения селективно нейтральных мутаций, предполагающая ступенчатый характер их возникновения. Рисунок построен с учетом динамики нейтральных мутаций, рассмотренной в предыдущих разделах книги. В частности, учитывается следующее: если p — частота нейтрального аллеля в популяции в данный момент, то вероятность того, что до элиминации или фиксации этот аллель достигнет более высокой частоты x , будет равна p/x . Что касается характера изменения электрофоретической подвижности в результате мутирования, то предполагается, что полоса в геле, отвечающая этому аллелю, в результате мутации перемещается на один шаг в положительном или отрицательном направлении. На рис. 8.7 эти изменения отмечены знаками плюс или минус. Говоря более конкретно, предполагается, что аспарагиновая и глутаминовая кислоты обладают кислотными свойствами и заряжены отрицательно, а лизин и аргинин — щелочными и заряжены положительно; остальные же аминокислоты электрически нейтральны. Предполагается также, что электрофоретическая подвижность полностью определяется суммар-

ным зарядом белковой молекулы. Далее принимается, что один положительный заряд нейтрализует один отрицательный. В различные моменты времени (абсцисса) последовательности расположения в геле часто встречающихся (С) и редких (R) аллелей могут быть представлены в виде схем CC, CR, RCR и т. д. Видно, что в отличие от последовательности RCR появление таких последовательностей, как CRR, RRC и CRC, маловероятно. В частности, последовательность CRC предполагает наличие между двумя «частыми» аллелями, разделенными по крайней мере двумя мутационными событиями, одного «редкого»; ясно, что при эффективном числе аллелей, равном 1,2 (т. е. $4N_e v = 0,2$), такое событие крайне маловероятно.

Таким образом, рис. 8.7 ясно показывает, что вывод Балмера о равной вероятности встречаемости всех последовательностей аллелей ни на чем не основан, а его наблюдения никоим образом не противоречат теории нейтральности. Предположение о нейтрализации одного положительного заряда одним отрицательным Кимура и Ота [259] считают излишним упрощением, поскольку с помощью более чувствительных методов можно разделить два варианта, имеющие заряды одного знака, но разной величины. В таком случае как исключение из правила Балмера возможно появление последовательности CRR.

В рамках модели ступенчатого мутирования уровень гетерозиготности должен быть ниже, чем в том случае, когда выявляется каждое мутационное изменение. Используя модель ступенчатого мутирования, представленную на рис. 8.6, и рассматривая панмиктическую популяцию эффективной численности N_e , Ота и Кимура [457] показали, что средняя гомозиготность, или ожидаемая величина суммы квадратов аллельных частот, в условиях равновесия равна

$$\bar{H}_0 = E \left[\sum_i x_i^2 \right] = 1/\sqrt{1 + 8N_e v}, \quad (8.75)$$

где x_i — частота аллеля A_i , а v — скорость мутирования на locus на поколение. В более общем случае, согласно этим авторам, ожидаемая величина произведения частот аллелей, отстоящих на k шагов друг от друга ($k=0, 1, 2, \dots$), равна

$$C_k = E \left[\sum_i x_i x_{i+k} \right] = \bar{H}_0 \lambda_1^k, \quad (8.76)$$

где

$$\lambda_1 = \frac{1 + 4N_e v - \sqrt{1 + 8N_e v}}{4N_e v}. \quad (8.76a)$$

Ожидаемая величина эффективного числа $\bar{n}_e$ аллелей в популяции, определенная как $1/\bar{H}_0$, равна

$$\bar{n}_e = \sqrt{1 + 8N_e v}. \quad (8.77)$$

Интересно сравнить это выражение с формулой (8.29) для модели бесконечного числа аллелей, которая предполагает, что каждая мутация является новой и может быть выявлена. Хотя при малых $4N_e v$ обе модели дают близкие значения n_e , при больших $4N_e v$ значение n_e , получаемое в рамках ступенчатой модели, гораздо меньше, чем в модели бесконечного числа аллелей. Например, если $4N_e v = 0,2$, то $n_e = \sqrt{1 + 8N_e v} = 1,18$, что очень близко к $n_e = 1 + 4N_e v = 1,20$. Однако при $4N_e v = 400$ величина n_e в рамках ступенчатой модели мутирования равна 28,3, а в модели бесконечного числа аллелей — 401.

После появления работы Оты и Кимуры [457] было опубликовано много статей, рассматривающих различные аспекты модели ступенчатого мутирования. Эта модель получила разные названия: модель состояний зарядов (charge-state model) [329, 98]; модель лестницы зарядов (charge-ladder model) [479]; модель Оты — Кимуры [272, 307], просто лестничная модель (ladder model) и др.

Относительно адекватности модели в смысле описания ею электрофоретически выявляемой изменчивости высказываются самые разные мнения. Некоторые авторы, например Джонсон [197], считают, что с биохимической точки зрения она совершенно не соответствует реальности. Рамшо и др. [479], основываясь на экспериментах с использованием вариантов гемоглобина, тоже пришли к выводу о полной неадекватности данной модели, поскольку с помощью электрофореза при разных значениях pH и разных концентрациях геля можно выявить значительную часть замещений аминокислот. Фуэрст и Феррелл [123], проведя аналогичное исследование вариантов гемоглобина, показали, что у 30—40% изученных вариантов наблюдались значительные отклонения от подвижности, соответствующей модели ступенчатого мутирования, хотя электрофоретическая подвижность молекулы определяется главным образом ее суммарным электрическим зарядом. Эти авторы считают, что модель бесконечного числа аллелей Кимуры и Кроу [246] или «смешанная модель» Ли [307] больше подходят для анализа электрофоретических данных, полученных в случае природных популяций.

С другой стороны, данные Ричардсона и др. [484] для *Drosophila mulleri* указывают на справедливость «лестничной модели» со «ступеньками», расположенными на одинаковом расстоянии друг от друга, а Маршалл и Браун [329] считают, что эту модель можно использовать в качестве первого приближения для предсказания степени и природы электрофоретически выявляемой изменчивости в природных популяциях. За недостатком места мы не можем и дальше рассматривать все аргументы «за» и «против» модели ступенчатого мутирования как способа исследования природы электрофоретически выявляемых вариан-

Тов. Читателям, интересующимся этим вопросом, мы советуем обратиться к содержательному обзору Брауна и др. [31]. Ниже я кратко изложу результаты, полученные при математическом обосновании модели ступенчатого мутирования.

Моран [368, 369] разработал весьма эффективный метод определения момента второго порядка аллельных частот и моментов более высоких порядков. Полученная им формула для дисперсии гомозиготности (или гетерозиготности) оказалась очень сложной, однако затем он вывел простую приближенную формулу

$$\text{Var}(H_0) = \theta / (3 + 11,25\theta + 13\theta^2 + 1,7\theta^3), \quad (8.78)$$

где $\theta = 4N_e v$. Моран [368] рассмотрел также ожидаемую величину дисперсии числа «шагов» в профиле частот:

$$\sigma^2 = E \left[\sum_i i^2 x_i^2 - \left(\sum_i i x_i \right)^2 \right]. \quad (8.79)$$

Хотя среднее профиля частот случайным образом блуждает во времени в пространстве аллелей, дисперсия σ^2 в условиях равновесия остается постоянной. Моран показал, что в равновесии

$$\sigma^2 = 2N_e v. \quad (8.80)$$

Например, если $4N_e v = 0,2$, то $\sigma^2 = 0,1$, а стандартное отклонение числа шагов $\sigma = 0,32$.

Найти распределение аллельных частот («спектра частот») в случае модели ступенчатого мутирования крайне трудно, и решить эту задачу до конца пока не удалось. Однако Кимура и Ота [262, 264] разработали теорию, позволяющую получить приближенное решение, достаточно точное для нашего случая, т. е. для $4N_e v \leq 1$. Не углубляясь в математические детали, я упомяну только одно важное свойство, отличающее распределение аллельных частот в рамках модели ступенчатого мутирования [обозначим соответствующую функцию через $\Phi_{\text{ступ}}(x)$] от распределения для обычной модели бесконечного числа аллелей [$\Phi(x)$ в (8.17)]: ожидаемое число аллелей, имеющих низкую частоту, в модели ступенчатого мутирования намного меньше. Например, при $4N_e v = 0,2$ $\Phi_{\text{ступ}}(0,001) = 111,3$, а $\Phi(0,001) = 200,1$. При $4N_e v = 1,0$ $\Phi_{\text{ступ}}(0,001) = 147,7$, тогда как $\Phi(0,001) = 1000$. На это различие моделей ранее указала Ота [438].

Еще более интересным свойством двух упомянутых моделей, связанным с предыдущим, является то, что фактическое число аллелей в этих двух случаях может быть совершенно различным даже при том, что эффективное их число ($\bar{n}_e = 1/\bar{H}_0$) примерно одинаково. Такая ситуация имеет место, когда параметр $4N_e v$ очень мал, а размер популяции N , напротив, очень велик (предполагается, что величина $4N_e v$ для обеих моделей одинакова).

Обозначим через $\bar{n}_{\text{ф;ступ}}$ ожидаемое фактическое число аллелей в популяции; тогда

$$\bar{n}_{\text{ф;ступ}}(N) = \int_{1/2N}^1 \Phi_{\text{ступ}}(x) dx. \quad (8.81)$$

При $4N_e v = \text{const}$ эта величина очень медленно увеличивается с увеличением N . Действительно, Кимура и Ота [264], основываясь на разработанной ими приближенной теории, получили следующий результат: $\bar{n}_{\text{ф}}(\infty) = 4,09$ для $4N_e v = 1$ (в данном случае эффективное число аллелей равно 1,73). С другой стороны, в рамках обычной модели бесконечного числа аллелей $\bar{n}_{\text{ф}}$ бесконечно возрастает с увеличением N при постоянном $4N_e v$. На быстрое увеличение $\bar{n}_{\text{ф}}$ с ростом N указывает тот факт, что уже в выборке размером 250 особей, взятой из популяции, для которой $4N_e v = 1$, число различных аллелей равно 6,79 (см. табл. 1 в работе [97]). Отметим, что эффективное число аллелей $\bar{n}_e = 4N_e v + 1 = 2,0$ не слишком отличается от соответствующей величины $\bar{n}_e = 1,73$ для модели ступенчатого мутирования.

Позднее Кестен [216], основываясь на модели Оты — Кимуры, разработал эффективный математический аппарат для определения числа различных аллелей в большой выборке. Согласно этому автору, асимптотическое значение $\Lambda(n, N)$ (число различных аллелей в выборке размера n из популяции, состоящей из N размножающихся особей) не зависит от скорости мутирования. Кестен показал, что для больших n и N ($n \rightarrow \infty$, $N \rightarrow \infty$ при $2N \geq n$)

$$\Lambda(n, N) \longrightarrow h^{-1}(p_2^+ + p_1^-) \lambda(n), \quad (8.82)$$

где $h^{-1}(p_2^+ + p_1^-)$ — постоянная, которая в простейшем случае, представленном на рис. 8.6, равна 2. Весьма интересными свойствами обладает функция $\lambda(n)$. Приведем некоторые ее значения: для $3 \leq n < 16$ $\lambda(n) = 1$; для $16 \leq n < 3814280$ $\lambda(n) = 2$; для $3814280 \leq n < 10^{1656520}$ $\lambda(n) = 3$. Легко видеть, что $\lambda(n)$ растет медленнее, чем $\log \log \dots \log n$. Однако при $n \rightarrow \infty$ $\lambda(n)$ бесконечно возрастает, так что — в противоположность утверждению Кимуры и Оты [264] — $\bar{n}_{\text{ф}}(\infty) = \infty$. Хотя аппроксимацию Кимуры и Оты нельзя считать удовлетворительной со строго математической точки зрения, она тем не менее вполне разумна. Действительно, реальный размер выборки не может быть больше 10^{100} , поскольку, как известно, даже общее число пар протон — электрон в нашей Вселенной предполагается равным $\sim 10^{80}$. Трудно представить себе выборку, размер которой превышает $n = 10^4$. Модель ступенчатого мутирования исследовалась также Кинг-

мэном [271, 272] в связи с тем, что она представляет интерес с чисто математической точки зрения.

Руководствуясь более конкретными задачами, Уэйр и др. [577] разработали статистическую теорию для проверки соответствия данных по электрофоретически выявляемому полиморфизму белков модели ступенчатого мутирования в случае нейтральности. Авторы использовали эту теорию для анализа данных Айалы и Трейси [12] по 27 полиморфным локусам в шести популяциях *Drosophila willistoni* и получили следующие результаты. Из 27 изученных локусов примерно для 15% (около $1/4$ всех полиморфных локусов) с достоверностью и во всех случаях наблюдалось несоответствие предсказаниям модели; для следующих 15% локусов в одних популяциях обнаруживалось достоверное отклонение от этих предсказаний, а в других — соответствие им. По остальным локусам наблюдалось достаточно хорошее согласие с моделью. В первых двух случаях показано, что аллели, характеризующиеся экстремальными значениями электрофоретической подвижности, испытывали направленное против них давление отбора.

Маруяма и Кимура [341] исследовали уровень генетической изменчивости в конечной популяции, предполагая, что как нейтральные, так и слабо вредные мутации возникают по «ступенчатому» типу. Используя два примера, в которых мутации с очень слабым вредным эффектом возникают в 100 раз чаще, чем совершенно нейтральные, а их селективная невыгодность только в 10 раз превышает суммарную скорость мутирования, эти авторы показали, что эффективное число аллелей возрастает с увеличением размера популяции гораздо медленнее, чем в случае только полностью нейтральных мутаций.

Такахата [540] рассмотрел сложную модель ступенчатого мутирования, в которой аллели, кодирующие белок с определенным зарядом («электроморф» по терминологии Кинга и Оты [269]), могут находиться в K состояниях (т. е. каждый электроморф порождается K аллельными состояниями). В этой модели предполагается, что мутации возникают со скоростью v_a для электрофоретически выявляемых и со скоростью v_m для электрофоретически невыявляемых («молчащих») аллелей. В первом случае при изменении состояния аллеля последний перемещается с одинаковой частотой на один шаг либо в положительном, либо в отрицательном направлении. В случае «молчащих» замен аллель переходит с одинаковой частотой в любое из оставшихся ($K-1$) состояний в том же классе заряда. Таким образом, скорость превращения аллеля в какой-либо другой, электрофоретически отличный от него аллель равна $v_a/(2K)$, а в электрофоретически идентичный аллель — $v_m/(K-1)$. Предполагая селективную нейтральность аллелей, Такахата показал, что в состоянии равновесия средняя гомозиготность с учетом как вы-

являемых, так и невыявляемых методом электрофореза аллелей равна

$$\bar{H}_0 = \frac{1}{K} \left[\frac{1}{\sqrt{(1+2\theta_s)}} + \frac{(K-1)^2}{K(1+\theta_T) - (1-\theta_s)} \right], \quad (8.83)$$

где $\theta_T = 4N_e v_T$, $\theta_s = 4N_e v_s$, $v_T = v_s + v_m$. Средняя гомозиготность при учете только электрофоретически выявляемых аллелей равна

$$C_0 = 1/\sqrt{(1+2\theta_s)}. \quad (8.84)$$

Эта формула эквивалентна формуле (8.75) и представляет собой частный случай соотношения (8.83) для $K=1$. С другой стороны, при $\theta_s=0$ уравнение (8.83) сводится к $\bar{H}_0 = (K-1+\theta_T)/(K-1+K\theta_T)$, что эквивалентно (8.33а).

Стимулом к созданию модели Такахаты послужило сообщение Сингха и др. [516], которые, проводя электрофорез при четырех разных условиях и используя тест на устойчивость к нагреванию, обнаружили 37 аллелей локуса ксантиндегидрогеназы *Drosophila pseudoobscura*, в то время как ранее выявлялось только 6 аллелей этого локуса. Таким образом, модификация условий анализа привела к выявлению в 6 раз более высокой степени полиморфизма, чем при помощи стандартной методики электрофореза. Согласно Сингху и др. [516], средняя гетерозиготность по локусу *Xdh*, определенная стандартным методом и с помощью всех пяти методов, составила 0,44 и 0,73 соответственно. Применяя модель Такахаты для анализа этих данных, для которых $C_0=0,56$, $\bar{H}_0=0,27$ и $K=5$, получаем $\theta_s=1,09$ и $\theta_T=3,47$. Отсюда отношение θ_s/θ_T , представляющее собой долю мутаций, приводящих к появлению электрофоретически различных аллелей, равно 0,32. Такахата также отметил, что при малом θ_s и $K \geq 2$ полная генетическая изменчивость в популяции ближе к величине, предсказываемой моделью бесконечного числа аллелей («модели Кимуры — Кроу»), чем к величине, полученной с помощью модели ступенчатого мутирования.

8.6. Модель бесконечного числа сайтов

Эта модель была впервые сформулирована Кимурой [232]; основанием для ее построения послужили соображения, состоящие в том, что общее число нуклеотидных сайтов в геноме человека очень велико — $(3 \sim 4) \cdot 10^9$ пар оснований, тогда как скорость мутирования на сайт должна быть очень малой величиной — вероятно, порядка 10^{-8} — 10^{-9} на поколение. Предполагается, что полное число сайтов, по которым могут происходить мутации, так велико, а скорость мутирования на сайт так мала, что любая новая мутация происходит в новом, ранее гомоаллельном сайте, т. е. в сайте, по которому не было расщепления.

Поскольку мутация, возникшая в популяции конечного размера, за конечный отрезок времени либо фиксируется, либо элиминируется, число расщепляющихся сайтов в каждый данный момент остается конечным, даже если число имеющихся сайтов бесконечно велико, а в популяцию в каждом поколении привносится какое-то число новых мутаций.

С использованием этой модели была разработана математическая теория [232], позволяющая определить число гетерозиготных нуклеотидных сайтов, поддерживающихся в конечной популяции в условиях постоянного «притока» мутаций. Например, можно показать, что если в популяции некоего млекопитающего эффективной численности 10 000 особей селективно нейтральные мутации происходят со скоростью, равной двум заменам нуклеотидов на гамету на поколение, то среднее число гетерозиготных нуклеотидных сайтов на особь, порождаемых этими мутациями, равно примерно $8 \cdot 10^4$.

Эта модель, впоследствии названная «моделью бесконечного числа сайтов» [234], оказалась очень полезной для анализа проблем эволюции и изменчивости на молекулярном уровне. Не углубляясь в детали математической стороны дела, я суммирую далее основные результаты, касающиеся числа гетерозиготных расщепляющихся сайтов, числа сегрегирующих сайтов и распределения частот мутаций в популяции [231].

Рассмотрим панмиктическую диплоидную популяцию эффективного размера N_e и предположим, что в каждом поколении новые мутации возникают в v_m сайтах. Пусть s — селективное преимущество данного мутантного аллеля по отношению к уже имеющемуся в популяции аллелю, так что селективное преимущество гомозигот и гетерозигот составляет $2s$ и s соответственно (случай «полудоминантности» по приспособленности). Обозначим начальную частоту данной мутации через p . Предположим, что параметры s и p одинаковы для мутаций по различным сайтам. Можно показать, что в условиях равновесия, когда процессы появления мутаций и их случайной элиминации уравновешивают друг друга, число гетерозиготных сайтов на особь равно

$$H(p) = (2v_m/s) [u(p) - p], \quad (8.85)$$

где $u(p) = (1 - e^{-Sp}) / (1 - e^{-s})$, а $S = 4N_e s$. $u(p)$ — вероятность того, что в конце концов мутация фиксируется в популяции. Если мутации селективно нейтральны ($s \rightarrow 0$), то (8.85) сводится к

$$H(p) = 4N_e v_m p (1 - p). \quad (8.86)$$

Если популяция состоит из N особей и данная мутация уникальна (вполне разумное предположение для молекулярных мутаций), то $p = 1/(2N)$. Тогда число гетерозиготных сайтов на особь равно $H(1/2N)$.

В частном случае селективно нейтральных мутаций из (8.86) имеем

$$H(1/2N) \approx 2v_m N_e / N. \quad (8.87)$$

Пусть скорость мутирования на гамету на поколение равна $v_{\text{гам}}$. Тогда

$$v_m = 2N v_{\text{гам}}, \quad (8.88)$$

и уравнение (8.87) можно представить в виде

$$H(1/2N) \approx 4N_e v_{\text{гам}}. \quad (8.89)$$

В дальнейшем, когда мы ограничимся рассмотрением лишь определенного участка генома (например, кодирующего участка, состоящего из 1000 нуклеотидных сайтов), а не всего генома, мы будем использовать вместо $v_{\text{гам}}$ просто v .

В разд. 8.4 тот же символ $H(p)$ использовался нами для обозначения ожидаемой суммы по всем поколениям частот гетерозигот, порождаемых данной мутацией, до ее фиксации или элиминации при условии, что начальная частота ее равна p . Было показано [см. уравнение (8.57)], что в случае селективно нейтральной мутации, появляющейся в популяции впервые, данная величина равна $2N_e/N$. Поскольку в популяции возникает $2N v_{\text{гам}}$ разных новых мутаций на поколение, ожидаемое число гетерозиготных сайтов в условиях равновесия равно $(2N_e/N) \cdot 2N v_{\text{гам}} = 4N_e v_{\text{гам}}$, что соответствует уравнению (8.89).

В качестве примера рассмотрим популяцию человека, в которой селективно нейтральные мутации возникают в среднем со скоростью 10^{-8} на сайт на поколение. Если геном (гаплоидный набор хромосом) состоит из $3,5 \cdot 10^9$ пар нуклеотидов, то $v_{\text{гам}} = 3,5 \cdot 10^9 \cdot 10^{-8} = 35$. Если эффективный размер популяции человека за время его эволюции был равен $N_e = 10^5$, то из (8.89) мы будем иметь $H(1/2N) = 4 \cdot 10^5 \cdot 35 = 1,4 \cdot 10^7$, т. е. число гетерозиготных нуклеотидных сайтов, порождаемых нейтральными мутациями, равно 14 млн. Хотя это и большая величина, она составляет всего 0,4% общего числа нуклеотидных сайтов генома.

Рассмотрим более общий случай, учитывающий естественный отбор. Из уравнения (8.85) мы получаем

$$H(1/2N) = 4N \left(\frac{v_{\text{гам}}}{s} \right) \left[\frac{1 - \exp(-Sp)}{1 - \exp(-S)} - \frac{1}{2N} \right], \quad (8.90)$$

где $S = 4N_e s$. При $|s| \ll 1$ (8.90) сводится к

$$H(1/2N) \approx 4N_e \left(\frac{v_{\text{гам}}}{s} \right) \left[\frac{2s}{1 - \exp(-4N_e s)} - \frac{1}{2N_e} \right]. \quad (8.91)$$

Если каждая мутация обладает явным селективным преимуществом, так что $4N_e s \gg 1$, то эта формула еще более упрощается:

$$H(1/2N) \approx 8N_e v_{\text{гам}}, \quad (8.92)$$

откуда видно, что мутации с явным селективным преимуществом вносят в гетерозиготность вклад, вдвое больший, чем селективно нейтральные мутации [см. (8.89)], при условии, что оба типа мутаций возникают с одинаковой частотой. С другой стороны, если мутация явно неблагоприятна, так что $4N_e s' \gg 1$, где $s' = -s$, то (8.91) сводится к

$$H(1/2N) \approx 2v_{\text{гам}}/s'. \quad (8.93)$$

Ота [435, 437] высказала предположение, что большинство полиморфных вариантов белков порождается очень слабо вредными мутациями и что белковый полиморфизм в популяциях поддерживается благодаря равновесию между мутационным процессом и отбором, когда интенсивность отбора не слишком сильно превышает скорость мутирования. Подобная мысль высказывалась также Райтом [604].

Вообще говоря, я полагаю, что мутации с явно благоприятным или явно вредным эффектом не могут вносить большого вклада в гетерозиготность особи из-за редкой встречаемости благоприятных и быстрой элиминации вредных мутаций.

Посмотрим теперь, чему равно среднее число расщепляющихся сайтов $I_1(p)$ в популяции [примем $p = 1/(2N)$]. Если каждый мутант обладает явным селективным преимуществом, так что $4N_e s \gg 1$, то

$$I_1(1/2N) \approx 4N_e v_{\text{гам}} [\ln(8NN_e s) + \gamma + 1], \quad (8.94)$$

где $\gamma = 0,5772 \dots$ — константа Эйлера. Если же каждый мутант явно невыгоден, так что $4N_e s' \gg 1$ (где $s' = -s$), то

$$I_1(1/2N) \approx 4N_e v_{\text{гам}} [-\ln(2N_e s'/N) - \gamma + 1]. \quad (8.95)$$

В случае селективно нейтральных мутаций среднее число расщепляющихся сайтов равно

$$I_1(1/2N) \approx 4N_e v_{\text{гам}} [-\ln(2N) + 1]. \quad (8.96)$$

Например, для популяции с $N_e = 10^5$ и $N = 10^6$ при скорости возникновения нейтральных мутаций 10^{-8} на сайт на поколение и общем числе пар нуклеотидов в геноме $3,5 \cdot 10^9$ получим $I_1(1/2N) = 2,2 \cdot 10^8$. Эта величина составляет примерно 6% от всех сайтов, хотя частоты мутаций в сегрегирующих сайтах обычно малы.

Наконец, рассмотрим распределение аллельных частот в условиях равновесия. Обозначим через $\Phi(p, x)$ ожидаемое число сайтов, частоты мутаций по которым находятся в интервале от x до $x+dx$. Если принять, что мутация по каждому отдельному

сайту уникальна, и обозначить $\Phi(1/2N, x)$ как $\Phi(x)$, то

$$\Phi(x) = \frac{2v_{\text{гам}} \int_x^1 G(x) dx}{V_{\delta x} G(x) \int_0^1 G(x) dx}, \quad (8.97)$$

где $G(\cdot)$ задается формулой (8.56в), а $M_{\delta x}$ и $V_{\delta x}$ — среднее и дисперсия изменения частоты данной мутации (x) за одно поколение. Эта формула, полученная Кимурой [228], распространяет результат Райта на случай необратимых мутаций [591, 593, 595]. Отметим, что она совпадает с формулой (8.73), за исключением множителя $2Nv_{\text{гам}}$, представляющего собой число распределенных по всей популяции сайтов, в которых в каждом поколении возникают мутации.

Более строгий и подробный анализ числа расщепляющихся сайтов был проведен Уоттерсоном [571] с использованием модели бесконечного числа сайтов (см. также [99]).

Хотя первоначально модель бесконечного числа сайтов была построена с учетом всего генома, она может применяться для отдельного генного локуса или цистрона (участка ДНК, кодирующего какой-то белок), поскольку каждый цистрон состоит из большого числа (порядка 1000) нуклеотидных сайтов. Ли [309] проанализировал распределение различий в нуклеотидном составе двух случайно выбранных цистронов в конечной популяции, используя модель бесконечного числа сайтов в ее применении к цистрону.

С помощью модели бесконечного числа сайтов Ота и Кимура [455] исследовали неравновесие по сцеплению (т. е. неслучайную ассоциацию мутантных форм) между двумя расщепляющимися нуклеотидными сайтами для популяции конечного размера. Основываясь на полученных результатах, они пришли к выводу, что между расщепляющимися нуклеотидными сайтами одного цистрона должна существовать высокая степень неравновесия по сцеплению.

8.7. Модель эффективно нейтральных мутаций

Обсуждая в гл. 7 взаимосвязь между скоростью эволюции и селективными ограничениями, мы часто для удобства разделяли мутации на два класса: полностью нейтральные мутации и мутации с явно выраженным вредным эффектом. При этом мы считали, что доля нейтральных мутаций уменьшается с увеличением функциональных ограничений. Однако в действительности между этими двумя типами мутаций, по-видимому, нет резкой границы, и не исключено, что мутации, занимающие промежу-

точное положение по отношению к этим двум экстремальным типам, имеют важное значение. Другими словами, как подчеркивала Ота [435—437], на молекулярном уровне значительная часть вновь возникающих мутаций может обладать очень слабым неблагоприятным эффектом. В популяции малого размера такие мутации ведут себя как селективно нейтральные, однако в больших популяциях они несомненно подвержены давлению отбора.

Впервые популяционно-генетические последствия наличия мутаций с очень слабым вредным эффектом исследовала Ота [442]. Предположив, что коэффициенты отбора против новых мутаций распределены по экспоненциальному закону, она показала, что скорость эволюции на поколение, выраженная через число мутационных замен, обратно пропорциональна эффективному размеру популяции данного вида.

С точки зрения теории нейтральности модель Оты имеет следующий недостаток: она полагает слишком малым число мутаций, которые в условиях большого размера популяции ведут себя как полностью нейтральные. Чтобы устранить этот недостаток, я предложил модель [238, 239], в которой для коэффициентов отбора мутаций было принято гамма-распределение. Эта модель была названа моделью эффективно нейтральных мутаций. Она основывается на предположении, что селективная нейтральность — это крайний случай селективной невыгодности мутаций, когда она становится бесконечно малой [261]. Более конкретно эта модель предполагает, что частотное распределение селективной невыгодности (обозначенной s') мутаций в разных сайтах одного гена (цистрона) соответствует гамма-распределению

$$f(s') = \alpha^\beta e^{-\alpha s'} s'^{\beta-1} / \Gamma(\beta), \quad (8.98)$$

где $\alpha = \beta / \bar{s}'$ ($\bar{s}'$ — средняя селективная невыгодность), а β — параметр, такой, что $0 < \beta \leq 1$. Если мы будем выражать селективный отбор через мальтузианский параметр [107], s' будет изменяться в пределах от 0 до ∞ . С другой стороны, если пользоваться обычным коэффициентом отбора, истинная область изменения s' будет ограничена интервалом (0, 1). Впрочем, поскольку мы предполагаем, что величина $\bar{s}'$ мала (обычно она составляет 10^{-3}), $f(s')$ при $s' < 0,1$ пренебрежимо мало, так что мы можем, не совершая большой ошибки, принять в качестве области интегрирования всю положительную часть оси.

В такой постановке мы не принимаем во внимание благоприятных мутаций, ограничиваясь рассмотрением вредных и нейтральных мутаций. Конечно, это является слишком большим упрощением, однако, как будет показано позже, модель, предполагающая, что благоприятные мутации также возникают с постоянной скоростью независимо от изменений окружающей сре-

ды, приводит к результатам, не соответствующим действительности.

Рассмотрим диплоидную популяцию эффективного размера N_e . Применим модель бесконечного числа сайтов (разд. 8.6) по отношению к локусу или цистрону, обозначив через v_e скорость

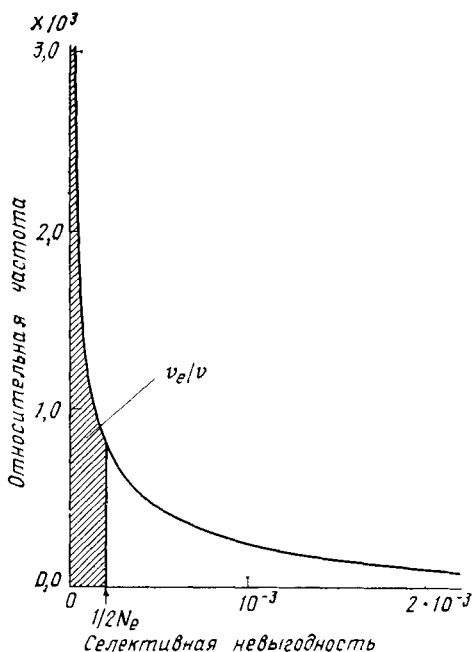


Рис. 8.8. Частотное распределение коэффициентов отбора для мутаций в разных сайтах одного цистрона (гена). Заштрихованная площадь отвечает доле эффективно нейтральных мутаций (здесь $\beta=0,5$; $\bar{s}=0,001$) [238].

мутирования для эффективно нейтральных мутаций, которая определяется следующим образом:

$$v_e = v \int_0^{1/2N_e} f(s') ds', \quad (8.99)$$

где v — скорость мутирования на цистрон. Другими словами, будем считать мутацию эффективно нейтральной, если интенсивность отбора против нее меньше, чем $1/(2N_e)$. При $2N_e \bar{s}' \gg 1$ (8.99) аппроксимируется выражением

$$v_e = \frac{v}{\Gamma(1+\beta)} \left(\frac{\beta}{2N_e \bar{s}'} \right)^\beta. \quad (8.100)$$

На рис. 8.8 приведено распределение $f(s')$ для случая $\beta=0,5$ и $\bar{s}'=10^{-3}$. Заштрихованная площадь соответствует доле эффективно нейтральных мутаций v_e/v при эффективном размере популяции $N_e=2500$. Эта величина уменьшается при увеличении размера популяции. Отметим, что, даже когда частота полностью нейтральных мутаций (для которых $\bar{s}'=0$) равна нулю, в рамках данной модели большой процент мутаций может быть эффективно нейтральным, если β мало. Отметим также, что при $0 < \beta < 1$ $f(0) = \infty$. Мы можем рассматривать величину β как меру уровня физиологического гомеостаза, а $\bar{s}'$ — как меру функциональных ограничений, налагаемых на молекулу. В предельном случае $\beta \rightarrow 0$ все мутации становятся нейтральными, а при $\beta = 1$ данная модель сводится к модели Оты, для которой при $2N_e\bar{s}' \gg 1$ $v_e/v = 1/(2N_e\bar{s}')$.

Скорость эволюции, выраженная через число мутационных замен на поколение и обозначенная нами k_g , равна

$$k_g = \int_0^{\infty} 2Nu f(s') ds', \quad (8.101)$$

где индекс «g» означает, что время выражается числом поколений, а не годами. В этой формуле u — вероятность того, что мутант, испытывающий направленное против него давление отбора интенсивности s' , в конце концов зафиксировается в популяции. Эта вероятность определяется как

$$u = [1 - e^{2s'(N_e/N)}] / (1 - e^{4N_e s'}) \quad (8.102)$$

(эта формула следует из (3.11), если подставить в нее $p=1/2N$ и $s=-s'$).

Тогда можно показать, что в грубом приближении

$$k_g \approx v_e, \quad (8.103)$$

т. е. скорость эволюции на поколение равна скорости возникновения эффективно нейтральных мутаций. Последнее утверждение можно считать обобщением хорошо известного соотношения $k=v$, справедливого для полностью нейтральных мутаций [см. (3.20)]. В случае $\beta=0,5$, представленном на рис. 8.8, скорость эволюции на поколение обратно пропорциональна $\sqrt{N_e}$ [при условии, что $N_e\bar{s}'$ велико; это видно из формулы (8.100)].

Пусть $\bar{H}_n$ — среднее число гетерозиготных сайтов на цистрон, так что

$$\bar{H}_n = \int_0^{\infty} H(1/2N) f(s') ds', \quad (8.104)$$

где $H(1/2N)$ задается формулой (8.90), в которой $v_{\text{гам}}=v$ и $s=-s'$. Для нахождения $\bar{H}_n$ удобна следующая аппроксимация:

$$\bar{H}_n = 8N_e v R^\beta \left[\frac{(1+R)^{1-\beta} + (1,5+R)^{1-\beta} - R^{1-\beta}}{1-\beta} - (1+R)^{-\beta} - \frac{(1,5+R)^{2-\beta} - (0,5+R)^{2-\beta}}{(1-\beta)(2-\beta)} \right], \quad (8.105)$$

где $R=\beta(4N_e \bar{s}')$, а $0 < \beta < 1$. В предельных случаях, когда либо $\beta \rightarrow 0$, либо $\bar{s}' \rightarrow 0$, эта формула сводится к $\bar{H}_n = 4N_e v$, что согла-

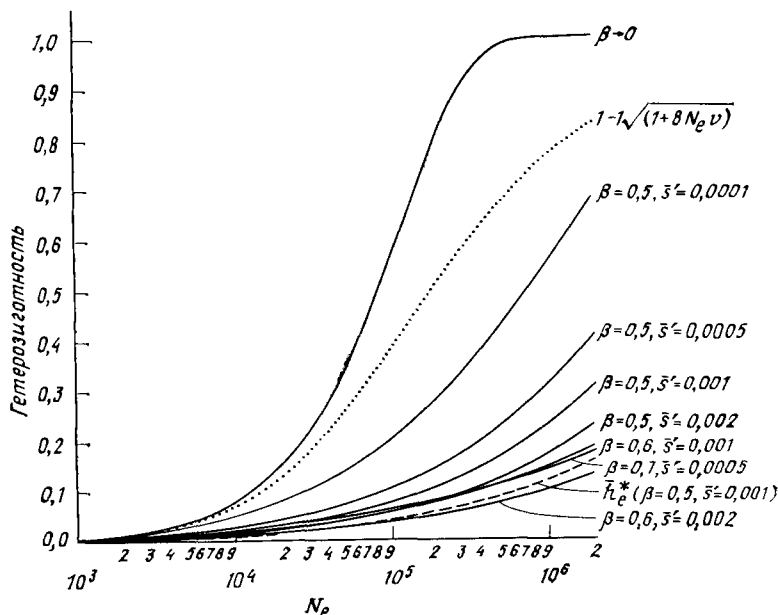


Рис. 8.9. Зависимость средней гетерозиготности на локус от эффективной численности популяции при различных значениях параметров β и $\bar{s}'$; скорость мутирования на локус v предполагается равной $2 \cdot 10^{-6}$ [238].

суется с уравнением (8.89), полученным в предыдущем разделе для полностью нейтральных мутаций.

Пусть $\bar{h}_{\text{ож}}$ — ожидаемая гетерозиготность рассматриваемого гена. В предположении независимости разных сайтов этого гена друг от друга

$$\bar{h}_{\text{ож}} = 1 - e^{-H_n}. \quad (8.106)$$

Это есть вероятность того, что данный ген гетерозиготен по крайней мере по одному сайту. Формула (8.106) предполагает независимое расщепление сайтов в цистроне и поэтому, строго говоря, неверна, однако для малых $\bar{H}_n$ она является достаточно

точной. На рис. 8.9 представлена зависимость ожидаемой гетерозиготности от N_e для различных значений параметров β и $\bar{s}'$ при $v = 2 \cdot 10^{-6}$. Отметим, что по сравнению с ситуацией, когда все мутации нейтральны ($\beta \rightarrow 0$), такой случай, как $\beta = 0,5$ и $\bar{s}' = 0,001$, представляет интерес из-за медленной скорости увеличения гетерозиготности при увеличении N_e . Пунктирная кривая на этом рисунке отвечает изменению гетерозиготности $\bar{h}_{\text{ож}} = 1 - 1/(1 + 8N_e v)^{1/2}$ в случае модели ступенчатого мутирования (см. разд. 8.5), а штриховая кривая — изменению $\bar{h}_e^* = 1 - 1/(1 + 8N_e v_e)^{1/2}$, где v_e — скорость возникновения эффективно нейтральных мутаций при $\beta = 0,5$ и $\bar{s}' = 0,001$. Во всех этих случаях скорость мутирования предполагается равной $2 \cdot 10^{-6}$.

В приведенных выше рассуждениях мы предполагали, что все мутации являются или вредными, или в лучшем случае нейтральными. Однако для полноты анализа необходимо рассмотреть эволюционный процесс и при условии, что некоторые мутации благоприятны. Итак, предположим, что благоприятные мутации возникают со скоростью v_A на поколение, а для коэффициента отбора s этих мутаций характерно гамма-распределение со средней $\bar{s}$ и параметром γ :

$$f_A(s) = \alpha^\gamma e^{-\alpha s} s^{\gamma-1} / \Gamma(\gamma), \quad (8.107)$$

где $\alpha = \gamma/\bar{s}$ и $\gamma = 0$.

Заметив, что вероятность закрепления отдельной мутации, имеющей коэффициент отбора s (> 0), равна $u = 2s(N_e/N) / (1 - e^{-4N_e s})$ [(3.12)], получим скорость эволюции, обусловленную благоприятными мутациями:

$$k_g = \int_0^\infty 2N v_A u f_A(s) ds = 4N_e v_A \bar{s} R^{\gamma+1} \sum_{j=0}^\infty (j+R)^{-\gamma-1}, \quad (8.108)$$

где $R = \gamma/(4N_e \bar{s})$. Это выражение можно аппроксимировать как

$$k_g = v_A \{4N_e \bar{s} + [\gamma/(6N_e \bar{s} + \gamma)]^\gamma\}. \quad (8.109)$$

При $N_e \bar{s} \gg 1$ имеем $k_g \approx 4N_e \bar{s} v_A$. Это означает, что, в полном противоречии с действительностью, скорость эволюции в очень большой популяции может быть очень большой, поскольку k_g прямо пропорционально N_e .

Вспомним теперь о сделанном нами вначале предположении, что возникновением полезных мутаций можно пренебречь и что частота вредных мутаций следует гамма-распределению (8.98). В рамках этих предположений мы показали, что скорость эволюции на поколение k_g уменьшается с увеличением N_e . Особый интерес представляет случай, когда $\beta = 0,5$ — промежуточному значению между $\beta = 0$ (все мутации нейтральны) и $\beta = 1$ (модель

экспоненциального распределения Оты). Ситуация, когда $\beta=0,5$, представляется более реалистичной, чем две другие.

Как указывалось выше, при $\beta=0,5$ скорость эволюции обратно пропорциональна $\sqrt{N_e}$, т. е.

$$k_g \propto 1/\sqrt{N_e}, \quad (8.110)$$

если $4N_e s' \gg 1$.

Таким образом, скорость эволюции за год (обозначим ее через k_1) равна

$$k_1 = k_g/g \propto 1/(g \sqrt{N_e}), \quad (8.111)$$

где g — продолжительность поколения в годах. Если, кроме того, продолжительность поколения обратно пропорциональна корню квадратному из эффективного размера популяции различных организмов, т. е.

$$g \propto 1/\sqrt{N_e}, \quad (8.112)$$

то, согласно (8.111), имеем

$$k_1 = \text{const}, \quad (8.113)$$

т. е. скорость эволюции за год является величиной постоянной, если скорость мутирования на поколение v постоянна у различных организмов.

Как подробно обсуждалось в гл. 4 (разд. 4.2), существуют убедительные свидетельства в пользу того, что для каждого данного белка скорость эволюции, выраженная через скорость замены аминокислот, примерно одинакова у разных организмов в расчете на год (но не на поколение). Эта закономерность особенно очевидна в случае α -гемоглобина млекопитающих.

Одно из критических замечаний, высказываемых в адрес теории нейтральности, связано с тем, что наблюдаемое постоянство скорости эволюции в расчете на год, а не на поколение противоречит предсказаниям этой теории. Дело в том, что скорость мутирования на поколение (но не на год) примерно одинакова у животных с совершенно разными продолжительностями поколения. Например, наблюдаемые скорости мутирования в случае летальных и видимых мутаций на локус на поколение примерно одинаковы у дрозофилы, мыши и человека. Если механизмы возникновения нейтральных мутаций и мутаций другого типа не отличаются друг от друга коренным образом, частота нейтральных мутаций на поколение должна быть приблизительно постоянной, и, исходя из формулы (3.20) $h=v$, скорость эволюции на поколение должна быть постоянной, что противоречит действительности.

Должен признаться, что это противоречие трудно разрешить в рамках теории нейтральности, хотя в случае «селекционист-

ской» теории сделать это еще труднее. Конечно, может быть справедливым сделанное ранее допущение (см., например, [256]), что скорость мутирования в расчете на год для нейтральных аллелей постоянна, но тогда возникает более серьезный вопрос: почему скорость возникновения нейтральных и почти нейтральных мутаций пропорциональна времени в годах? Этот вопрос особенно уместен в связи с данными по псевдогенам (см. разд. 7.4), в случае которых мутации, бывшие ранее вредными, в дальнейшем превратились в нейтральные.

Мнение о том, что спонтанные мутации в основном возникают за счет ошибок копирования в ходе репликации ДНК, представляется почти бесспорным [282]. Попытки оценить для видимых мутаций скорости мутирования на локус на один цикл репликации у *Drosophila*, мыши и человека дали величины $(2 \sim 50) \cdot 10^{-8}$, $4 \cdot 10^{-8}$ и $(2 \sim 20) \cdot 10^{-8}$ соответственно. Эти оценки получены путем деления скорости мутирования на локус на гамету на число генераций половых клеток, которое предполагается равным 11, 16 и 20 для дрозофилы, мыши и человека соответственно. Учитывая огромную разницу между временем генерации у этих организмов, мы должны сделать вывод, что скорости возникновения спонтанных мутаций у них сравнимы в расчете именно на поколение, а не на год.

Предположение о том, что эффективный размер популяции данного вида связан обратной зависимостью с временем генерации, вероятно, соответствует действительности, поскольку для крупных организмов (например, слона) обычно характерна большая продолжительность поколения и в то же время меньшее число особей на вид, чем для организмов малых размеров (например, для мыши). Однако, выполняется ли хотя бы приближенно зависимость $g \propto 1/\sqrt{N_e}$, пока неизвестно. В качестве примера сопоставим слона и мышь. Время генерации у слона в 40 раз больше, чем у мыши: у слона оно, вероятно, составляет около 20 лет, а у мыши — полгода. Если для этих животных выполняется зависимость $g \propto 1/\sqrt{N_e}$, то эффективный размер популяции мыши должен превышать размер популяции слона примерно в 1600 раз, что не является таким уж нереалистичным допущением. Заметим, что здесь имеется в виду эффективный размер всей популяции данного вида или подвида, а не локальной популяции (дема).

Исключением из закономерности эволюции белков, состоящей в постоянстве ее скорости в расчете на год, является медленная скорость молекулярной эволюции высших приматов. Это отклонение можно объяснить, предположив, что в ходе эволюции высших приматов увеличение g не сопровождалось соответствующим уменьшением N_e , в результате чего произведение $g\sqrt{N_e}$ в уравнении (8.111) стало значительно больше, чем в случае дру-

гих млекопитающих. Это привело к снижению скорости эволюции в расчете на год. Отметим, впрочем, что, как указывалось в гл. 4, еще существуют разногласия по поводу того, действительно ли у высших приматов замедлена скорость молекулярной эволюции, поэтому данное обсуждение является сугубо предварительным.

Вообще говоря, вполне возможно, что параметр β у насекомых больше, чем у млекопитающих, из-за более высокого уровня физиологического гомеостаза последних. Кондо [282] высказал предположение, что у высших форм (таких, как млекопитающие), обладающих совершенным гомеостазом, нейтральными оказывается большее число мутаций. Низкий уровень физиологического гомеостаза и частое вымирание отдельных колоний, вероятно, являются главными причинами того, что виртуальная гетерозиготность у организмов с очень большой численностью популяции (неотропические виды *Drosophila* [11] и *Escherichia coli* [357, 507]) не очень велика. В следующей главе (разд. 9.2) этот вопрос будет рассмотрен более подробно.

Модель эффективно нейтральных мутаций [238], которая обсуждается в настоящем разделе, является воплощением попытки сделать теорию нейтральности более строгой и адекватной. В этой модели предполагается, что молекулярная эволюция и полиморфизм в основном обусловлены случайным дрейфом мутаций с очень слабым неблагоприятным эффектом, т. е. фактически нейтральных мутаций. В этом отношении модель эффективно нейтральных мутаций напоминает теорию Оты о слабо вредных мутациях [435, 437, 440]. Однако между этими моделями есть существенные различия. Ота [440] утверждает, что в очень больших популяциях устойчивое равновесие между мутационным процессом и отбором устанавливается при достижении популяцией верхнего предела гетерозиготности, когда молекулярная эволюция прекращается или по крайней мере замедляется. В этих условиях закрепление мутаций происходит почти исключительно в результате резкого уменьшения размеров популяции (эффект «горлышка бутылки») во время видообразования. Я же, напротив, предполагаю, что даже в очень больших популяциях аллели, имеющие промежуточную частоту (что часто встречается у *Drosophila*; см. [506], табл. II), возникли как нейтральные мутации, подхваченные затем случайным дрейфом, и что эволюция, обусловленная дрейфом, у этих видов вряд ли прекращается.

И наконец, последняя биологическая проблема, которую мы должны рассмотреть. В рамках данной модели у каждого вида происходит постоянное накопление эффективно нейтральных, но на самом деле очень слабо неблагоприятных мутаций. Интенсивность отбора против таких мутаций (в смысле выживаемости особей и их воспроизведения, т. е. дарвиновской приспособлен-

ности), вероятно, составляет $\sim 10^{-5}$ или менее, однако при наличии в геноме 10^4 локусов, кодирующих белки, каждый из которых накапливает мутации со скоростью 10^{-6} на поколение, скорость уменьшения приспособленности может составить 10^{-7} на поколение. Представляет ли уменьшение приспособленности, происходящее с такой малой скоростью, угрозу выживанию и процветанию вида (не особи!) — вопрос спорный; впрочем, положение легко исправляется благодаря адаптивным генным заменам, которые должны происходить время от времени, скажем, раз в несколько сотен поколений (см. также разд. 6.7, где обсуждается случайное закрепление слабо вредных мутаций при стабилизирующем отборе).

8.8. Генетическое расстояние

Для оценки степени генетической дивергенции между популяциями (видами, расами и т. д.) по данным о генетической изменчивости необходимо найти подходящую меру расстояния между популяциями. В своей пионерской попытке построить филогенетическое древо рас человека на основании данных по частотам групп крови Кавалли-Сфорца и Эдвардс [43] при измерении генетического расстояния использовали не сами генные частоты, а корни квадратные из них. Этот способ измерения был ранее предложен Кавалли-Сфорца и Контерио [42], которые анализировали распределение генов, детерминирующих группы крови человека в популяции долины реки Парма. Их метод основан на следующих соображениях. Пусть $A_1, A_2, \dots, A_m$ — это m аллелей данного локуса; возьмем m -мерную декартову систему координат, в которой каждая ось представляет собой квадратный корень из частоты одного из аллелей. Тогда генетический состав популяции образует гиперсферу с радиусом, равным 1. (В действительности нам потребуется только $1/2^m$ -я часть этого пространства, где все координаты неотрицательны.) Генетическое расстояние между двумя популяциями, аллельные частоты в которых есть $(p_1, p_2, \dots, p_m)$ и $(p'_1, p'_2, \dots, p'_m)$, равно θ , причем

$$\cos \theta = \sum_{i=1}^m \sqrt{p_i} \sqrt{p'_i}. \quad (8.114)$$

В геометрической интерпретации θ — это угол (в радианах) между двумя векторами $(\sqrt{p_1}, \sqrt{p_2}, \dots, \sqrt{p_m})$ и $(\sqrt{p'_1}, \sqrt{p'_2}, \dots, \sqrt{p'_m})$, отражающими генетический состав двух данных популяций. Преимущество этого способа измерения генетического расстояния состоит в том, что в случае, когда генные частоты изменяются только под действием генетического дрейфа, ожидаемая величина дисперсии изменения θ на поколение является

постоянной, зависящей лишь от эффективного размера популяции (но не от частот генов). Поскольку $\theta = \pi/2$ соответствует полной замене генов, удобно использовать величину $2\theta/\pi$; тогда единицей расстояния будет одна генная замена.

Однако для указанной меры расстояния метод максимального правдоподобия трудноприменим из-за аналитических сложностей, поэтому Кавалли-Сфорца и Эдвардс использовали более удобную меру $(2\sqrt{2/\pi})\sqrt{1 - \cos \theta}$, которая представляет собой длину хорды, соединяющей две точки (популяции). Такая мера удобна, если дивергенция генных частот между популяциями отражает непосредственное действие случайного генетического дрейфа в течение сравнительно короткого периода времени и если влияние мутационного процесса пренебрежимо мало.

Однако при анализе эволюционных изменений за большой промежуток времени, когда важное значение играет кумулятивный эффект мутаций, необходима другая мера генетического расстояния. По-видимому, самая лучшая мера такого рода, используемая наиболее широко, предложена Неем [400]. Несомненно, популярность метода Нея объясняется простотой формулировки, легкостью применения и ясностью биологического смысла, в чем мы сейчас убедимся.

Рассмотрим две популяции (1 и 2) и предположим, что частоты аллеля A_i локуса A в популяциях 1 и 2 равны соответственно x_i и y_i ($i=1, 2, \dots$). Пусть $j_1 = \sum x_i^2$, $j_2 = \sum y_i^2$, а $j_{12} = \sum x_i y_i$, где j_1 — вероятность (коэффициент) идентичности двух случайно выбранных гомологичных генов в популяции 1, j_2 — соответствующая вероятность для популяции 2, а j_{12} — вероятность идентичности двух гомологичных генов, один из которых случайно выбран из популяции 1, а другой — из популяции 2. Эти вероятности определяются для каждого локуса, а затем усредняются по всем соответствующим локусам, для которых известны генные частоты. Пусть $J_1 = \bar{j}_1$, $J_2 = \bar{j}_2$, $J_{12} = \bar{j}_{12}$ будут этими средними величинами. Тогда «стандартное генетическое расстояние» по Нею равно

$$D = -\ln \frac{J_{12}}{\sqrt{J_1} \sqrt{J_2}}. \quad (8.115)$$

Это выражение можно представить как $D = -\ln I$, где

$$I = J_{12} / \sqrt{(J_1 J_2)}. \quad (8.116)$$

Величина I является нормированным коэффициентом идентичности генов.

Стандартное генетическое расстояние D должно служить мерой среднего числа различий в кодонах; в условиях полной изоляции оно дает суммарное число генных замен на локус, накопившихся после того, как рассматриваемые популяции диверги-

ровали от общей предковой популяции. Тот факт, что скорость молекулярной эволюции, выраженная числом мутационных замен, примерно постоянна в расчете на год (см. гл. 4), делает этот способ измерения генетического расстояния особенно полезным для исследования молекулярной эволюции.

Пусть v — скорость мутирования на локус для нейтральных аллелей; предположим, что эта скорость одинакова для всех локусов. Тогда ожидаемая идентичность в t -м поколении равна

$$E(I_t) = e^{-2vt}I_0, \quad (8.117)$$

где I_0 — вероятность идентичности (или гомозиготность) в предковой популяции ($t=0$). Эта формула непосредственно вытекает из следующего рассуждения: если мы используем модель бесконечного числа аллелей, которая предполагает, что каждая мутация приводит к появлению нового, не существовавшего ранее в популяции аллеля (см. разд. 8.2), и если две популяции изолированы друг от друга, вероятность идентичности двух генов из этих двух популяций уменьшается со скоростью $2v$ на поколение, так что

$$E(I_t) = (1 - 2v) E(I_{t-1}) = (1 - 2v)^t I_0 \approx e^{-2vt} I_0. \quad (8.118)$$

Это значит, что ожидаемое стандартное расстояние равно

$$E(D) \approx -2vt,$$

если рассматриваемая популяция находится в состоянии равновесия, и $J_1 = J_2 = J_0$. С другой стороны, если изоляция между популяциями не является абсолютной, как во многих локальных популяциях данного вида, то миграция будет препятствовать процессу дивергенции и генетическое расстояние не сможет стать очень большим (коэффициент идентичности субпопуляций в условиях миграции и возникновения мутаций анализировался Неем и Фельдманом [409]). Отметим, что, используя в качестве меры суммарного различия числа кодонов величину D , в общем случае необходимо оценивать ее, используя большое число локусов, случайно выбранных из генома, причем в это число должны входить как мономорфные, так и полиморфные локусы.

Скорость мутирования v для нейтральных аллелей может изменяться от локуса к локусу, о чем мы говорили в связи с распределением аллельных частот и средней гетерозиготности по многим локусам (см. разд. 8.2). Согласно работе [408], распределение скорости мутирования на локус хорошо аппроксимируется гамма-функцией, а полученный эмпирическим путем коэффициент вариации $\alpha = \bar{v}^2 / V_v$ примерно равен 1. Тогда соответствующая формула для вычисления генетического расстояния имеет вид

$$D_v = (1 - I)/I. \quad (8.120)$$

Вывод этой формулы приведен в работах [404, 405].

Идентичность двух видов по одному локусу ($j_{12}/\sqrt{j_1j_2}$) существенно различается для разных локусов. Обычно у позвоночных эта величина имеет U-образное распределение и часто равна 1 или 0 [404]. Это легко понять, заметив, что аллели данного локуса в ходе эволюции, как правило, либо закрепляются, либо теряются, если (как утверждает теория нейтральности) главную роль в определении их частот играет случайный дрейф генов. Следовательно, для определения истинного генетического расстояния D между видами необходимо проанализировать большое число генов. Сходная ситуация возникает при определении гетерозиготности данного вида на локус (H). Дисперсии этих оценок были проанализированы в работах [404, 405, 418] путем их разложения на две компоненты, одна из которых обусловлена межлокусной изменчивостью (межлокусная дисперсия), а вторая — выборкой генов в пределах одного локуса. Оказалось, что межлокусная дисперсия гораздо больше внутрилокусной, за исключением случая, когда величина n (число особей в выборке) крайне мала. Таким образом, для определения средней гетерозиготности и истинного генетического расстояния следует рассматривать не большое число особей на локус, а большое число локусов (см. также [313]).

Используя литературные данные, Ней [401] определил стандартные генетические расстояния между расами, видами, родами и т. д. различных организмов. Его результаты показывают, что генетическое расстояние между расами обычно равно 0,00~0,05, между подвидами — 0,02~0,20, между видами — 0,1~2,0, а генетическое расстояние между родами превышает 1.

Недавно Ней [405] опубликовал обзорную работу по определению генетического расстояния в связи с проблемами эволюции человека. Я остановлюсь на некоторых из полученных им интересных результатов. Расовые различия для человека на уровне белков соответствуют различиям между локальными расами у других организмов. Среднее число различий в кодонах между человеком и шимпанзе составляет около 1/19 соответствующих различий между человеком и лошадью. Что касается основных рас человека — европеоидной, негроидной и монголоидной, — дивергенция между ними, по оценкам Ней, произошла впервые около 120 000 лет назад — негроидная группа отделилась от европеоидно-монголоидной. Затем последняя около 60 000 лет назад разделилась на европеоидную и монголоидную расы. Генетическое сходство между человеком и шимпанзе свидетельствует о том, что дивергенция между ними произошла 4~6 млн. лет назад, если справедливо предположение о постоянной скорости замены аминокислот (см. разд. 4.1). Для ознакомления со сходными исследованиями в азиатско-тихоокеанском регионе читатели могут обратиться к работам [459, 460].

Поддержание генетической изменчивости

9.1. Формулировка проблемы, факты и трактовка с позиций теории нейтральности

Каков механизм поддержания генетической изменчивости, проявляющейся в виде полиморфизма белков? Этот вопрос имеет первостепенное значение для современной популяционной генетики.

В предыдущей главе я рассмотрел некоторые модели, использующиеся в молекулярной популяционной генетике и разработанные на основе метода диффузионных уравнений. Именно таким образом в рамках теории нейтральности ищется подход к решению этой важной и сложной проблемы. Далее, взяв в качестве примера гипотетическое млекопитающее, я количественно показал, как можно получить достоверные оценки уровня генетической изменчивости и ее распределения, используя набор значений параметров, полученных независимым путем (например, скорость замен оснований, оцененную исходя из результатов исследования эволюции глобиновых псевдогенов; долю нейтральных мутаций от общего числа мутаций «среднего» белка, полученную из частоты редких электрофоретически различимых вариантов; средний размер генов, кодирующих ферменты; обоснованные предположения об эффективном размере популяций). При этом моей главной целью было показать, что полиморфизм белков и их эволюция тесно связаны между собой и что полиморфизм белков представляет собой лишь переходную фазу молекулярной эволюции [256].

Конечно, любая оценка или предположение относительно значений упомянутых выше параметров не могут считаться абсолютно достоверными. Тем не менее тот факт, что при этом получается непротиворечивая и достаточно полная картина, убедительно свидетельствует в пользу теории нейтральности.

В этой главе мы обсудим, как согласуется теория нейтральности с результатами многочисленных полевых наблюдений и разнообразных лабораторных экспериментов. Кроме того, мы рассмотрим, каких успехов достигла она в объяснении данных по генетической изменчивости на молекулярном уровне, если сопоставить ее с конкурентами — «селекционистскими» теориями.

Прежде чем приступить к подробному обсуждению молекулярного полиморфизма, я бы хотел упомянуть о двух типах данных, имеющих для такого обсуждения основополагающее значение. Это, во-первых, распределение средней гетерозиготности (на локус на особь) у различных организмов, а во-вторых — распределение частот аллелей данного ферментного локуса у родственных видов.

Распределение гетерозиготности установлено в настоящее время для множества самых разных организмов. На рис. 9.1

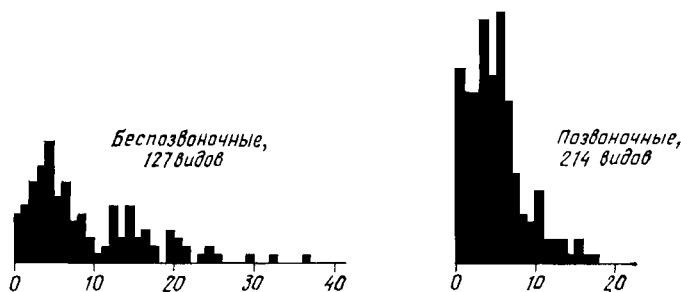


Рис. 9.1. Распределение средней гетерозиготности на локус для позвоночных и беспозвоночных (Ней, Граур, личное сообщение). По оси абсцисс отложена гетерозиготность, или, точнее, «виртуальная гетерозиготность» (т. е. единица минус сумма квадратов аллельных частот) на локус (в %), по оси ординат — число видов.

представлены такие распределения для позвоночных (214 видов) и беспозвоночных (127 видов). По оси абсцисс отложена гетерозиготность, или, точнее, единица минус сумма квадратов аллельных частот, на локус (в процентах). Согласно терминологии, принятой в данной книге, эта величина представляет собой виртуальную гетерозиготность (или, по терминологии Нея [401], «генное разнообразие»); для диплоидного вида со случайным скрещиванием она совпадает с гетерозиготностью в обычном смысле слова. Очень похожие гистограммы приведены на рис. 2 работы Нево [420], суммировавшего данные по 93 видам беспозвоночных и 135 видам позвоночных, у которых было исследовано по 14 и более локусов. Из этих рисунков видно, что средняя гетерозиготность редко превышает 30%, т. е. величина $\bar{H}_e$ в природных популяциях изменяется в пределах 0,0—0,3.

Здесь уместно вспомнить утверждения Левонтина [302], который в 1974 г., основываясь на имевшихся в то время данных по 12 видам (человек, мышь, *Peromyscus*, мечехвост и 8 видов дрозофилы), заявил, что наблюдаемый интервал изменения гетерозиготности у всех видов лежит в «чувствительной области» между 0,056 и 0,184. Это соответствует значениям $N_e u$ [т. е. про-

изведению эффективного размера популяции на скорость мутационного процесса; см. уравнение (8.15) предыдущей главы], лежащим в интервале от 0,015 до 0,057, т. е. крайние значения $N_e v$ отличаются друг от друга не более чем в четыре раза. Если учесть, что скорость мутирования v может быть сходной у различных организмов, то, как считает Левонтин, мы сразу получаем, что размеры популяций у всех многоклеточных очень близки. Далее он утверждает, что «заведомая абсурдность такого предположения — веский довод против нейтралистского объяснения наблюдаемой гетерозиготности».

Однако после того как были получены более полные данные, стало ясно, что средняя гетерозиготность некоторых организмов очень мала, а иногда даже равна нулю. В частности, очень низкая гетерозиготность характерна для крупных млекопитающих [38]. Приведу несколько примечательных примеров. При проведенном в Гренландии исследовании [3] 13 ферментных локусов у белого медведя электрофоретическая подвижность соответствующих ферментов была совершенно одинаковой у изученных 52 особей, т. е. все эти локусы оказались мономорфными ($H_e = 0$). Полиморфизм отсутствовал и у морского слона (*Mirounga angustirostris*) по 21 белку, кодируемому 24 локусами. Боннел и Силандер [26], которые провели это исследование, объясняют обнаруженный ими абсолютный мономорфизм фиксацией аллелей, происшедшей в результате истребления значительной части этого вида тюленей в прошлом веке. Однако, как было показано в конце разд. 8.1, быстрое уменьшение размера популяции в течение небольшого числа поколений вряд ли может привести к значительному уменьшению гетерозиготности. Следовательно, для данного вида действительно характерны низкие значения гетерозиготности. В рамках теории нейтральности это можно объяснить, предположив, что популяции крупных млекопитающих, как правило, имеют небольшой эффективный размер, поэтому значение $N_e v$ для них очень мало. Предположим, что скорость мутирования для нейтральных электроморфов $v_{0(s)} \approx 10^{-6}$ на поколение (т. е. в несколько раз выше соответствующей скорости для рассмотренного в гл. 8 гипотетического млекопитающего, продолжительность поколения у которого равна одному году). Тогда при $N_e < 1000$ ожидаемая средняя гетерозиготность будет меньше 0,004.

Следовательно, единственный серьезный вопрос, на который еще нужно ответить в связи с критикой Левонтина, — это почему средняя гетерозиготность редко превышает 0,3 даже у видов, популяции которых имеют огромную наблюдаемую численность. Согласно Айале [6], для тропической плодовой мушки *D. willis-toni* характерен чрезвычайно большой размер популяций с ареалом, простирающимся на несколько миллионов квадратных километров, а общее число особей в данный момент времени несомненно намного превышает 10^9 . Айала подверг критике тео-

рию нейтральной эволюции, заявив, что если принять в качестве оценки снизу значение $N_e = 10^9$ и предположить, что скорость мутирования $v = 10^{-7}$ [256], то мы получим эффективное число аллелей, равное $4N_e v + 1 = 401$. Другими словами, полученное исходя из предположения о нейтральности значение гетерозиготности будет равно $\bar{H}_e = 4N_e v / (4N_e v + 1)$, т. е. практически 100%, тогда как наблюдаемая величина средней гетерозиготности *D. willistoni* составляет всего около 18% (см. также [11]).

Ота и Кимура [457] в ответ на это критическое замечание указывают, что значение $N_e = 10^9$ может быть завышенной оценкой (если говорить не о числе размножающихся особей на поколение, а о числе особей, используемом для определения ожидаемой гетерозиготности). Во-первых, если размер популяции флуктуирует от поколения к поколению, это число определяется минимальным размером популяции [уравнение (3.5)]. А во-вторых, для того, чтобы установилось равновесное по гетерозиготности состояние после быстрого увеличения численности популяции при исходно низкой гетерозиготности, необходимо время порядка величины, обратной скорости мутирования (если $1/v \ll \ll N_e$), или размеру популяции (если $1/v \gg N_e$) [см. уравнение (8.13)]. Как предполагают Ота и Кимура [457], последствия резкого уменьшения размера популяции тропических видов дрозофилы в период последнего оледенения материков (около 30 000—10 000 лет назад) заметны в генетическом составе этих видов еще и сегодня. Согласно Колдеру [37], в ледниковый период на больших площадях тропических лесов преобладал сухой и прохладный климат. Весьма вероятно, что вследствие этого популяции тропических дрозофил были тогда гораздо меньше, чем теперь. Отсюда следует, что равновесное значение гетерозиготности, ожидаемое для условий теперешнего процветания вида, пока не достигнуто. Возможно, что относительная численность различных видов *Drosophila* довольно быстро изменяется (в масштабах времени, соответствующих эволюционным изменениям), поэтому корреляция между гетерозиготностью и нынешним размером популяций может быть очень небольшой.

Далее, как указывали Маруяма и Кимура [342], для *Drosophila* характерны локальное вымирание и реколонизация субпопуляций, что приводит к резкому уменьшению размера их численности (см. п. 5 в разд. 3.3). Если в уравнении (3.10) предположить, что скорость локального вымирания (λ) колоний значительно выше скорости обмена особями между субпопуляциями (m), то эффективный размер популяции N_e рассматриваемого вида станет гораздо меньше, чем произведение числа колоний (n) на эффективный размер отдельной колонии ($\tilde{N}$). Действительно, при малом m/λ из уравнения (3.10) получаем $N_e \approx \approx n\tilde{N}m/\lambda$. Не исключено, что для *D. willistoni* соответствующий

эффективный размер популяции на два порядка меньше, чем минимальная оценка, принятая Айалой. Еще одно замечание по поводу критики Айалы: значение $v=10^{-7}$, полученное Кимурой и Отой [256], — это скорость мутирования для нейтральных мутаций в расчете на год, а не на поколение. Дрозофилы, размножающиеся в тропических лесах в течение всего года, могут дать около 10 поколений в год, и в рамках теории нейтральности соответствующая скорость мутирования для нейтральных электроморфов $v_{0(г)}$ будет равна 10^{-8} (более серьезное обсуждение взаимосвязи между скоростью мутирования «эффективно нейтральных аллелей» и N_e см. в разд. 8.7). Тогда, исходя из значений $N_e=10^7$ и $v_{0(г)}=10^{-8}$, мы получим $H_e=0,29$, что все еще больше наблюдаемого значения гетерозиготности (0,177), однако достаточно близко к ней для того, чтобы аргументы Айалы в значительной мере потеряли свою убедительность.

На небольшой эффективный размер популяций многих видов *Drosophila* в ходе их эволюции указывает также тот факт, что эти виды отличаются друг от друга по ряду хромосомных перестроек (см., например, [86]). Поскольку хромосомные перестройки типа инверсий и транслокаций, как правило, являются вредными в гетерозиготном состоянии, а в гомозиготном имеют нормальную приспособленность, формально можно считать, что они проявляют отрицательное сверхдоминирование, поэтому для их закрепления в ходе эволюции необходим малый эффективный размер популяции. Прекрасный математический анализ соответствующих вопросов был проведен Ланде [289]. Отметив, что виды обычно бывают представлены множеством популяций (демов), которые флуктуируют по величине и подвергаются локальному вымиранию и реколонизации, Ланде вычислил вероятность фиксации U в деме вновь возникшей хромосомной перестройки. При этом скорость фиксации R задается выражением $R=2N_a\mu U$, где μ — скорость возникновения хромосомных перестроек на гамету, а N_a — число размножающихся особей в деме. Далее, если n — число частично изолированных популяций, составляющих некий вид, то вероятность того, что данная популяция (или ее потомки) распространится по всему виду, равна $1/n$. Таким образом, вероятность фиксации перестройки для вида в целом равна вероятности ее фиксации в отдельном деме, т. е. R .

Для многих организмов, характеризующихся высокой скоростью размножения, процессы случайной локальной элиминации и реколонизации могут преобладать над межпопуляционной миграцией, что, как упоминалось выше, приводит в ходе эволюции к резкому уменьшению эффективной численности популяции. В пионерском исследовании Слаткина [518] изучалось влияние локального вымирания и реколонизации на генетическую изменчивость вида. Затем эти процессы были более подробно рассмот-

рены в уже упоминавшейся работе Маруямы и Кимуры [342]. Я считаю, что указанные работы, как и исследование «эффекта горлышка бутылки» [415], должны приниматься во внимание при рассмотрении генетической изменчивости организмов с очень большой наблюдаемой численностью популяций. Эффективный размер популяции может уменьшаться и по другим причинам (например, из-за различия в числе размножающихся самцов и самок). Поэтому в общем случае величина N_e , которая фигурирует в формуле $\bar{H}_e = 4N_e v / (4N_e v + 1)$ для определения средней гетерозиготности, должна быть гораздо меньше числа особей, участвующих в данный момент в размножении. Как отмечал Ней [406], в каждом случае, когда удастся грубо оценить эффективный размер популяции, наблюдаемая гетерозиготность почти всегда оказывается ниже ожидаемой, однако это расхождение не составляет серьезной проблемы для теории нейтральности. Другими словами, с ее позиций не составляет большого труда объяснить наличие «крайне высокого уровня» генетической изменчивости (как это часто считается). Действительно, уровень генетической изменчивости, предсказанный на основе теории нейтральности, скорее слишком высок, поэтому привлекать другие механизмы для его поддержания нет необходимости. Мне кажется, здесь возникают проблемы как раз для селекционистских теорий, предполагающих различные причины поддержания в популяции аллельных вариантов (например, «маргинальное сверхдоминирование», обусловленное взаимодействиями генотип — среда в гетерогенной среде или частотно-зависимый отбор, при котором преимущество получают более редкие генотипы): эти механизмы приведут к дальнейшему увеличению степени гетерозиготности, если не пренебречь вкладом мутационного процесса.

Теперь вкратце рассмотрим второй тип данных, а именно — распределение аллельных частот локуса, кодирующего данный фермент, у различных видов. Подобные данные имеются для дрозофилы и других организмов. На рис. 9.2 [134] представлено распределение частот аллелей локуса *Xdh*, кодирующего ксантиндегидрогеназу (точнее — группы локусов *Xdh*, поскольку генетическая идентичность этого локуса у отдаленно родственных организмов не доказана). Распределение имеет U-образную форму. Такие же U-образные распределения характерны почти для всех локусов, кодирующих ферменты. Этот факт согласуется с тем, что распределение частот аллелей различных локусов в пределах вида также является U-образным (см. рис. 8.3, иллюстрирующий эту закономерность в случае камбалы). Обычно это распределение хорошо описывается (если не рассматривать избытка редких аллелей) уравнением (8.17), в котором параметр $4N_e v$ оценивается из наблюдаемой средней гетерозиготности. Это еще один довод в пользу теории нейтральности. С ее

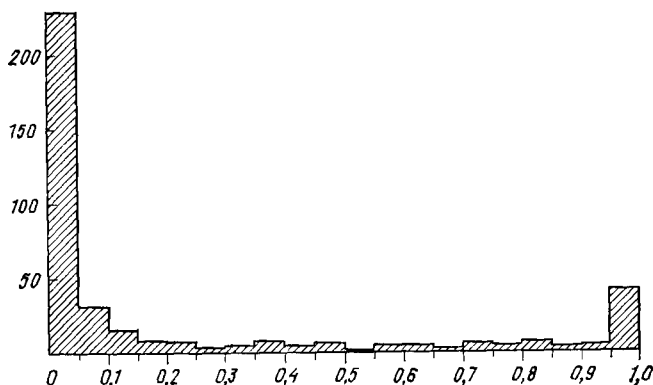


Рис. 9.2. Распределение частот аллелей локуса (точнее, группы локусов) *Xdh*, построенное Годжборн [134] по данным для различных организмов. По оси абсцисс отложена частота аллелей, по оси ординат — их число.

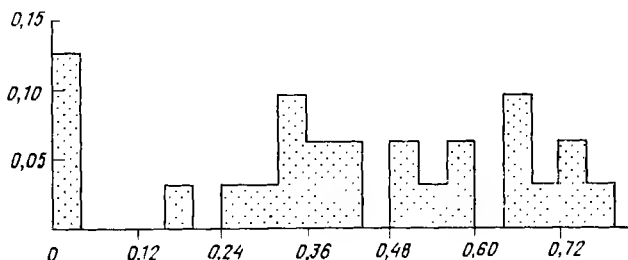


Рис. 9.3. Распределение гетерозиготности по локусу *Xdh* у различных видов *Drosophila* [134]. $H_e=0,345$. По оси абсцисс отложена гетерозиготность, по оси ординат — относительная частота.

позиций наблюдаемая частота аллеля у данного вида достигается в результате случайного дрейфа, а не благодаря адаптации к условиям среды. Другими словами, соответствующая частота достигается «случайно», а не «по необходимости». Следовательно, с течением времени эта частота случайным образом изменится, причем скорость такого изменения можно рассчитать. Эту точку зрения подтверждает также достаточно хорошее соответствие с теорией распределения генетических расстояний между различными видами (разд. 9.3).

Подобно распределению аллельных частот, мы можем рассмотреть распределение гетерозиготности по данному локусу у родственных видов. На рис. 9.3 приведено такое распределение для локуса *Xdh* различных видов *Drosophila*. Отметим, что гетерозиготность изменяется от 0 до почти 0,8. Сходное распределение представлено в работе [506] для 10 белков 28 видов *Drosophila* (см. табл. II из этой работы). Что касается локуса *Xdh*, то данная таблица свидетельствует о мономорфности по этому

ферменту у 32% видов, тогда как у 4% видов уровень гетерозиготности превышает 0,60, а среднее значение гетерозиготности равно 0,275 (небольшие различия результатов работ [134] и [506], очевидно, обусловлены ошибкой выборки). Похожая картина наблюдается и для других ферментных локусов. Тот факт, что уровень гетерозиготности по одним и тем же (гомологичным) локусам резко различается даже у разных видов в пределах рода *Drosophila*, изменяясь от мономорфизма до очень высокой степени полиморфизма, говорит о том, что простое равновесие между отбором и мутационным процессом [269] не может быть главной причиной полиморфизма ферментов — для этого необходим гораздо более динамичный процесс.

9.2. Экспериментальная проверка теории нейтральности

Являясь научной гипотезой, теория нейтральности полиморфизма белков должна быть проверена экспериментально. Аллели, кодирующие изозимы, можно исследовать на селективную нейтральность, например в опытах с «ящичной» популяцией дрозофилы, в которой два или более аллелей имеют разные частоты, или путем измерения компонентов приспособленности различных генотипов в контролируемых условиях. Таким образом, очевидно, можно ответить на вопрос, являются ли полиморфные формы белков селективно нейтральными или же они поддерживаются в популяции за счет балансирующего отбора. Нейтралистская позиция кажется особенно уязвимой для проверок такого рода.

Действительно, по этому поводу было опубликовано множество работ. Вышли в свет и продолжают появляться статьи, сообщающие о существовании балансирующего и других типов отбора и, таким образом, направленные против теории нейтральности. Оказалось, однако, что решить вопрос подобным способом гораздо труднее, чем представлялось вначале. На это существуют по крайней мере две основные причины. Во-первых, крайне трудно однозначно показать, что наблюдаемые селективные различия действительно обусловлены самими полиморфными аллелями, а не соседними с ними локусами, находящимися по отношению к изучаемому локусу в состоянии неравновесия по сцеплению. Во-вторых, нет гарантии, что условия проведения эксперимента действительно имитируют те условия среды, в которых данный вид существует в природе.

Из-за этих трудностей (в особенности из-за первой) большая часть результатов была признана неубедительной. Критический разбор относящихся к данному вопросу экспериментальных лабораторных данных проведен Левонтином [302] и Райтом [604]. Левонтин в своей работе ([302] с. 250—260) убедительно показывает, как легко получить в ходе лабораторных экспериментов

ложные закономерности, имитирующие сверхдоминирование и другие формы отбора, и как трудно выявить слабый отбор даже в том случае, когда его интенсивность составляет несколько процентов. Аналогичным образом Райт ([604], с. 298), внимательно проанализировав различные сообщения об отборе в лабораторных условиях, пришел к следующим выводам. Если эксперименты проводятся с использованием нескольких отдельных линий, в них обычно обнаруживаются резкие селективные различия между аллозимами, выражающиеся в виде преимущества редкого генотипа. Термин «аллозим» здесь обозначает формы ферментов, кодируемые разными аллелями. (Строго говоря, это определение относится к генетически контролируемым «изозимам», или множественным молекулярным формам ферментов. Однако поскольку эти два термина — «изозимы» и «аллозимы» — в популяционной генетике часто используются как синонимы, я тоже буду придерживаться этого). Дальнейшие исследования, однако, показали, что найденные селективные различия могут быть «отделены» от исследуемых аллозимов, что свидетельствует об их принадлежности к сцепленным с изучаемым локусом генам. В тщательно спланированных экспериментах, где генетический состав популяции хорошо гомогенизирован путем использования большого числа независимых по происхождению линий, селективные различия между аллозимами либо слабы, либо вовсе не обнаруживаются. По словам Райта, «эти результаты, впрочем, не доказывают полную нейтральность исследованных аллозимов; они свидетельствуют лишь о том, что дифференциальный отбор обычно имеет малую интенсивность (порядка 10^{-3} и меньше)».

Алкогольдегидрогеназа

Из всех изученных к настоящему времени полиморфных белков наиболее убедительные доказательства отбора имеются, пожалуй, для алкогольдегидрогеназы. Действительно, полиморфизм по этому локусу у *Drosophila melanogaster* интенсивно изучался Кларком и его сотрудниками с целью доказать, что отбор действует непосредственно именно на данный локус, а не на локусы, находящиеся с ним в неравновесии по сцеплению [51, 52]. Для выявления прямого действия естественного отбора Кларк использовал следующий подход. Прежде всего на основании результатов биохимического и физиологического анализа постулируется (и по возможности идентифицируется) соответствующий селективный агент и устанавливается его взаимосвязь с продуктом гена. Затем на основе этих данных формируются экспериментальные популяции, причем генные частоты подбираются таким образом, чтобы получить предсказуемые результаты. Кроме того, анализируются распределения частот генов и экологических факторов, с тем чтобы дать убедительное объяснение на-

блюдаемых закономерностей распределения генных частот в природных популяциях. Другими словами, проводится активный поиск корреляции между частотами генов или генотипов и определенными факторами среды. Согласно Кларку [51], эта стратегия прочно базируется на принципах экологической генетики, причем он оптимистически заявляет, что если в результате ее применения наличие естественного отбора будет показано даже только для семи-восьми локусов одного организма, «спор между «нейтралистами» и «селекционистами» будет наконец разрешен».

Локус *Adh* у *D. melanogaster* имеет два часто встречающихся полиморфных аллеля: *F* (fast, т. е. «быстрый» — имеется в виду электрофоретическая подвижность) и *S* (slow — «медленный»). Эти аллели часто обозначаются *Adh^F* и *Adh^S*. Согласно Кларку [51], при использовании в качестве субстрата этанола активность *F*-формы алкогольдегидрогеназы вдвое превышает активность *S*-формы. С другой стороны, *F*-форма менее устойчива к нагреванию. Селективным агентом в данном случае является спирт, и Кларк [51] предположил, что алкогольдегидрогеназа выполняет роль фермента детоксикации у *D. melanogaster*, которая в природных условиях живет в гниющих фруктах, т. е. «прямо-таки купается в этаноле и других спиртах». Это привело к предположению, что в среде с повышенной концентрацией этанола отбор будет идти в пользу аллеля *F*; вскоре это положение получило подтверждение в лаборатории Кларка (см. [370], где было также показано, что отбор в данном случае является частотно-зависимым). Данные, подтверждающие наличие отбора, были также получены Брискоу и др. [29], которые установили, что частота аллеля *Adh^F* у особей *D. melanogaster*, отловленных в винном погребе, выше, чем у тех, что были отловлены на расположенной поблизости мусорной свалке. Кроме того, экспериментально исследовав различия в смертности генотипов на среде, содержащей этанол (12,5%), эти авторы показали, что гомозигота *Adh^S/Adh^S* значительно менее устойчива к алкоголю, чем гетерозигота или гомозигота *Adh^F/Adh^F*. Томпсон и Кайзер [550], используя аллель *S*, который обозначается *Adh^S(ve)* и продукт которого имеет ферментативную активность, в два раза более низкую, чем продукт нормального аллеля *S*, исследовали относительную жизнеспособность от стадии яйца до вылета у особей, несущих этот аллель, а также аллели *Adh^F* и *Adh^S*, при наличии и в отсутствие этанола и н-бутанола. В их опытах наблюдалась положительная корреляция между выживаемостью и активностью фермента на среде с алкоголем.

В то время когда появились упомянутые статьи, доказательства наличия отбора по локусу *Adh* казались убедительными. Однако впоследствии в ходе более тщательно разработанных экспериментов были получены результаты, противоречащие или коренным образом отличающиеся от приведенных выше. Обшир-

ные исследования австралийских генетиков показали, что устойчивость *Drosophila melanogaster* к этанолу детерминруется большим числом генов, а не аллелями локуса *Adh*, причем различий в частоте аллеля *F* у особей, отловленных в австралийских винных погребах и вне их, обнаружено не было [322, 426]. Гибсон и др. [127] провели опыты по отбору на повышенную устойчивость к этанолу в двух популяциях *D. melanogaster*, начальная частота *Adh^F* в которых составляла 0,5. Полученные этими авторами результаты свидетельствуют о том, что успешный отбор на устойчивость к этанолу не обязательно сопровождается увеличением частоты аллеля *F* или повышением уровня активности алкогольдегидрогеназы. Авторы отмечают, что их результаты ставят под сомнение справедливость утверждения Кларка [51] и других генетиков, что различия в активности алкогольдегидрогеназы *D. melanogaster*, кодируемой аллелями *F* и *S*, имеют в природных популяциях адаптивное значение в отношении устойчивости к алкоголю; напротив, они соответствуют выводам работы [322] о том, что адаптация дрозофилы к повышенному содержанию алкоголя в винных погребах не зависит от системы *Adh*.

Позднее Окшотт и Гибсон [426] сообщили о результатах пяти экспериментов, поставленных ими с целью окончательного выяснения этого вопроса. Ни один из них не указывал на то, что наличие в среде этанола приводит к специфическому отбору особей генотипа *F/F* и *F/S* с высокой активностью фермента и элиминации особей *S/S*, для которых характерна низкая ферментативная активность. Напротив, в этих экспериментах выживаемость особей *S/S* на среде с этанолом была выше, причем некоторые данные свидетельствовали о том, что такое преимущество обусловлено низкой (а не высокой) активностью алкогольдегидрогеназы. Это можно объяснить тем, что ацетальдегид, образующийся из этанола под действием алкогольдегидрогеназы, весьма токсичен, вследствие чего высокая активность этого фермента не полезна, а вредна. Несмотря на очевидные натяжки, это рассуждение было привлечено для объяснения более высокой чувствительности к спирту японцев по сравнению с европейцами [528]. Дали и Кларк [71] недавно сообщили, что отношение числа выживающих мух генотипа *F/F* к *S/S* было выше на среде с 2%-ным этанолом, чем на дистиллированной воде. Авторы считают, что их экспериментальные результаты подтверждают положение о наличии отбора по локусу *Adh* в соответствии с работой [51]. Однако в свете рассмотренных выше результатов австралийских авторов утверждение Кларка [51] вызывает сильные сомнения.

Частотно-зависимый отбор и другие типы естественного отбора

По-видимому, впервые мысль о том, что в роли универсального механизма поддержания высокого уровня полиморфизма

ферментов в природных популяциях может выступать отбор, зависящий от частоты, высказали Кодзима и Ярбро [280]. Полученные ими экспериментальные данные показали, что в случае аллелей *F* и *S* локуса, кодирующего эстеразу-6 (*Est-6*) *D. melanogaster*, наблюдается значительное преимущество редких генотипов. Метод, использованный этими авторами для оценки относительной приспособленности трех генотипов, заключался в нахождении отношения наблюдаемой на стадии имаго частоты каждого генотипа к частоте, ожидаемой исходя из генных частот в предыдущем поколении. Данный метод, очевидно, не может дать правильных оценок приспособленности, когда генотипы различаются еще и по плодовитости. В связи с этим показано (см. [48, 476]), что при наличии отбора, происходящего после определения генных частот, оценки приспособленности создают ложное впечатление, будто отбор зависит от частоты, даже если реальные величины приспособленности постоянны. Более того, в большинстве случаев этот кажущийся отбор принимает вид преимущества редких генотипов. Для подкрепления своей гипотезы о наличии отбора, зависящего от частоты, Ярбро и Кодзима [616] провели эксперимент, в котором в течение 30 поколений измерялась частота аллелей *F* и *S* локуса *Est-6* в популяциях, выращиваемых на двух разных средах, в повторностях с двумя различными начальными частотами. В результате было обнаружено, что в трех популяционных ящиках частоты аллелей быстро приближались к частотам исходной популяции (30% *F* и 70% *S*), хотя в четвертой ящичной популяции, начальные частоты в которой составляли 10% *F* и 90% *S*, существенного отклонения от этих частот не наблюдалось.

С тех пор было проведено множество экспериментальных исследований с целью выяснить, существует ли для других локусов частотно-зависимый отбор типа постулированного Кодзимой и Ярбро. Изучая аллели *F* и *S* локуса *Adh* *D. melanogaster*, Кодзима и Тобари [279] нашли подтверждение существованию отбора, зависящего от частоты, с преимуществом редких генотипов. Для выяснения механизма такого отбора Хуанг и др. [187] исследовали локус *Est-6*, используя метод так называемой «кондиционированной среды». Этот метод состоит в изменении питательной среды личинками какого-либо генотипа, развивающимися в среде перед внесением в нее личинок того же или другого генотипа, выживаемость которого требуется определить. Авторы показали, что кондиционирование среды личинками данного генотипа уменьшает жизнеспособность личинок того же самого генотипа, развивающихся на этой среде, что предполагает истощение в ней запаса питательных веществ, необходимых для данного генотипа, либо выделение личинками в среду веществ, вредных для «своего» генотипа. Однако в поставленных позднее более тщательно разработанных экспериментах эффект

кондиционированной среды отсутствовал [87]. С другой стороны, Морганом [370] было ясно показано наличие частотно-зависимого отбора по обоим упомянутым локусам.

Для разрешения этой проблемы Иосимару и Мукаи [617] выполнили тщательно разработанные широкомасштабные эксперименты с использованием аллелей Adh^F и Adh^S *D. melanogaster*. Опыты выполнялись в два этапа. Сначала из выборки природной популяции Огасавара были выделены 563 вторые хромосомы, а затем на их основе получены 124 хромосомные линии, содержащие аллель F , и 120 линий, содержащих аллель S . Их хромосомы не несли инверсий и были выбраны случайным образом. Используя эти линии, провели серию из пяти экспериментов по конкуренции между генотипами F/F и S/S с начальными частотами аллеля F , равными 0,82; 0,67; 0,5; 0,33 и 0,17. Гомозиготные по аллелям F и S личинки первого возраста были помещены в различном количественном соотношении в пробирки, и затем определялась частота мух каждого генотипа. Относительная жизнеспособность исследованных генотипов не изменялась при изменении частоты гена. Средняя жизнеспособность особей S/S по отношению к жизнеспособности особей F/F составляла $1,0825 \pm 0,0144$ (в пробирках с 9 мл среды) и $1,0879 \pm 0,0127$ (в пробирках с 1,5 мл среды). Эти данные полностью противоречат результатам Моргана [370]. В экспериментах второго типа, где жизнеспособность определялась по методу Cy (с вариациями генных частот в пробирках при определении жизнеспособности особей, имеющих случайное сочетание двух вторых хромосом), доказательств наличия частотно-зависимого отбора получено не было. Эти результаты согласуются с выполненной ранее работой Ямадзакки [606], который в серии экспериментов всесторонне исследовал полиморфизм по сцепленному с полом локусу $Est-5$ (аллели F и S) *D. pseudoobscura*. Ямадзакки не обнаружил различий между генотипами по компонентам приспособленности и не получил данных в пользу существования частотно-зависимого отбора.

Разумеется, отрицательные результаты не являются доказательством того, что рассмотренные аллели совершенно нейтральны с точки зрения естественного отбора. Возможно, очень слабый отбор на них все же действует. Как правило, эксперименты по выявлению естественного отбора имеют небольшую чувствительность: интенсивность отбора, меньшую чем 0,01, с их помощью измерить уже невозможно. Однако Мукаи и др. [377], используя несущую инверсию хромосому $In(2L)t$ в изолированной популяции дрозофилы, смогли показать, что интенсивность отбора по ферментным локусам (если таковой отбор существует) не должна превышать величины порядка 10^{-4} . К такому выводу авторы пришли, проанализировав следующие результаты. В популяции Исигакидзима частота инверсии второй хромосомы

In(2L)t составляет около 60%. Три изоферментных локуса в хромосомах с инверсией почти мономорфны, однако в других (стандартных) хромосомах имеют высокую степень полиморфизма. Хромосомы, несущие инверсию, появились в популяции около 1000 поколений тому назад; эта величина была получена на основе известной скорости кроссинговера и низкого уровня полиморфизма. За это время частоты летальных и слабо вредных мутаций достигают равновесных значений, но для мутаций, на которые действует очень слабый положительный отбор, этого времени для достижения равновесия недостаточно. Вообще говоря, распределение жизнеспособностей гомозигот по инвертированной хромосоме неотличимо от такового для стандартной хромосомы. Предполагается, что в момент появления в популяции (т. е. 1000 поколений тому назад) хромосома *In(2L)t* была мономорфной. Этого времени недостаточно для достижения равновесия по полиморфизму. Если полиморфизм по этим локусам поддерживается за счет отбора в пользу гетерозигот, жизнеспособности гомо- и гетерозигот по стандартной хромосоме, высокополиморфной в отношении ферментных локусов, должны четко различаться, тогда как в случае почти мономорфной инвертированной хромосомы таких различий быть не должно. И в самом деле, эти различия отсутствовали — по крайней мере были не более 1%. В предположении, что имеется примерно 350 полиморфных локусов с 20%-ной гетерозиготностью каждый, мы получаем оценку коэффициента отбора против гомозигот на locus — 0,00013.

Для выявления отбора по ферментным локусам часто использовались ящичные популяции дрозофилы. Основанные на этих данных выводы о существовании балансирующего отбора (преимущество редкого генотипа или сверхдоминирование) делались многими авторами (например, [8, 118, 555]). Типичный результат таких экспериментов заключается в том, что частоты аллелей в ящичной популяции с любыми исходными частотами быстро приближаются к равновесному состоянию, характерному для природных популяций. Мукаи и Ямадзак [382] попытались оценить коэффициенты отбора, исходя из данных упомянутых выше работ. Оказалось, что полученные таким образом оценки довольно высоки: порядка 0,1, а иногда даже 0,2~0,4. Отметив, что коэффициенты отбора, равные или превышающие 10%, слишком высоки для обычных полиморфных аллелей изоферментного локуса, Мукаи и Ямадзак пришли к выводу, что данные экспериментальные результаты не отражают действия одного исследуемого гена.

Если говорить о предполагаемой распространенности частотно-зависимого отбора, то нередко высказывается мнение, что разные генотипы обычно адаптированы к слегка различающимся условиям среды или используют пищевые ресурсы данной экс-

логической ниши неодинаковым образом. В этом случае — если выполнен целый ряд дополнительных условий — устанавливается стабильный полиморфизм с преимуществом редкого генотипа. Было предложено несколько моделей поддержания подобного полиморфизма. Однако математический анализ [350] показал, что эти модели применимы только в том случае, когда относительный размер ниш изменяется в узких пределах.

Кларк и Аллендорф [53] предложили модель, по которой относительные скорости реакций, катализируемых двумя аллозимами, изменяются в зависимости от концентрации субстрата. Приняв целый ряд дополнительных допущений (приспособленность является монотонно возрастающей функцией скорости реакции и т. д.), они показали, что в этом случае будет иметь место частотно-зависимый отбор, и высказали предположение, что именно таким образом в значительной мере поддерживается полиморфизм ферментов (см. также [52]). С другой стороны, Мэйнард Смит и Хекстра [350] считают недостатком модели Кларка и Аллендорфа то, что она применима только в случае, когда полиморфизм может поддерживаться лишь в узком интервале изменения условий среды; по их мнению, из-за узости диапазона эта модель вряд ли может иметь широкое применение. Эти авторы также отметили, что в общем случае модели поддержания полиморфизма в гетерогенной среде требуют довольно точного соответствия приспособленности размеру ниши (если селективное преимущество локуса не очень велико), и в связи с этим подвергли серьезному сомнению правомочность привлечения данного механизма в качестве одной из основных причин широко распространенного полиморфизма.

Мне хотелось бы сделать здесь еще одно замечание. Часто утверждают, что частотно-зависимый отбор с преимуществом редкого генотипа не сопровождается возрастанием генетического груза. Однако это утверждение не всегда верно. Как показано в разд. 6.4, только при выполнении некоторого дополнительного условия [см. формулу (6.27) и ее обсуждение] три генотипа в точке равновесия действительно становятся эквивалентными и генетический груз не появляется. Во всех других случаях он неизбежно возникает, а уж в популяции конечной численности существует всегда.

Гаплоидные организмы

Одним из наиболее популярных механизмов, при помощи которого селекционисты объясняют широкую распространенность полиморфизма белков, является сверхдоминирование, или преимущество гетерозигот. Следовательно, изучение генетической изменчивости гаплоидных организмов представляет особый интерес для разрешения противоречий между нейтралистской и селекционистской теориями; у гаплоидных организмов гетерози-

готовность отсутствует, поэтому генетическая изменчивость не может поддерживаться за счет сверхдоминирования. Существует ли у гаплоидных организмов белковый полиморфизм?

Пионерская работа в этой области была выполнена Милкмэн [357] на *Escherichia coli*. Исследовав электрофоретическую изменчивость пяти ферментных локусов в 829 клонах *E. coli* из 156 выборок из различных природных источников, Милкмэн показал, что уровень генетической изменчивости *E. coli* не слишком сильно отличается от такового для многих диплоидных видов. Поскольку *E. coli* — гаплоидный организм, она не может быть гетерозиготной, и термин «гетерозиготность» в его обычном смысле здесь неприменим. Однако мы можем вычислить виртуальную гетерозиготность по каждому локусу, равную единице минус сумма квадратов аллельных частот. Эта величина является долей гетерозигот среди воображаемых зигот, каждая из которых получилась бы от случайного сочетания двух гаплоидных особей. Среднее значение виртуальной гетерозиготности для пяти локусов, вычисленное по данным табл. 1 из работы [357], составляет $\bar{H}_e = 0,242$, или около 24%. Эта величина превышает (хотя и не очень сильно) среднюю гетерозиготность типичного диплоидного организма. Полученные данные нанесли роковой удар гипотезе о сверхдоминировании как универсальной причине полиморфизма генов. Действительно, начиная с этого времени энтузиазм селекционистов по поводу сверхдоминирования постепенно спадал, и они отдавали все большее предпочтение частотно-зависимому отбору.

Отрицая гетерозис как универсальную причину генного полиморфизма, Милкмэн [357] в то же время утверждал, что его наблюдения не согласуются с теорией нейтральности. Его доводы были основаны на зависимости между ожидаемым и наблюдаемым значениями величины $\bar{n}_e - 1$, где $\bar{n}_e$ — среднее эффективное число аллелей на локус [см. формулу (8.29)]. Согласно Милкмэну (см. также [358]), размер популяции *E. coli* превышал 10^6 по крайней мере в течение $4 \cdot 10^{10}$ поколений, или $4 \cdot 10^7$ лет (эта бактерия обнаружена в организме большинства млекопитающих по всему земному шару). Поскольку скорость мутирования для нейтральных аллелей вряд ли может быть меньше 10^{-8} , эффективное число аллелей минус единица ($\bar{n}_e - 1$), равное, согласно теории нейтральности, $4N_e v$, должно составлять 400. С другой стороны, наблюдаемые значения величины $\bar{n}_e - 1$ для пяти локусов лежат в интервале от 0 до 1. Действительно, наблюдаемые значения $\bar{n}_e$ (обратные величины квадратов аллельных частот) приведены в табл. 1 статьи Милкмэна [357], и среднее по пяти локусам равно 1,38. Это значит, что наблюдаемое значение $\bar{n}_e - 1$ равно всего 0,38.

Генетическая изменчивость и структура популяций *E. coli* были исследованы не так давно Силандером и Левином [507], получившими новые данные, в значительно меньшей степени противоречащие теории нейтральности. Эти авторы проанализировали 20 ферментов 109 клонов *E. coli* из различных природных источников и получили значение виртуальной гетерозиготности, равное 0,4718, что вдвое превышает оценку Милкмэна [357]. Гораздо более важное (и уж несомненно очень интересное) открытие состояло в том, что число отдельных генотипов в популяции было довольно ограниченным: в 109 изученных клонах авторы обнаружили только 98 различных типов электроморфов. Идентичные пары клонов были выделены даже от хозяев, разделенных большими географическими расстояниями, причем они были выделены в разное время. Примечательный пример: клон, выделенный от ребенка в Массачусетской детской больнице, был неотличим электрофоретически от обычной линии *E. coli* K12 (возможность заражения здесь исключалась). Более того, в нескольких случаях два клона отличались друг от друга только по одному-двум локусам из 20 проанализированных. По мнению авторов, эти наблюдения легко объяснить, предположив, что в природных популяциях *E. coli* рекомбинация происходит крайне редко (что противоречит более раннему утверждению Милкмэна [357], считавшего, что у бактерий, обитающих в одном хозяине, рекомбинация происходит регулярно). Действительно, если бы аллели разных локусов рекомбинировали случайным образом, число различных комбинаций было бы астрономическим, и, по оценкам Силандера и Левина [507], даже частота комбинации с наивысшим ожиданием составила бы только $5,35 \cdot 10^{-6}$; тогда вероятность обнаружения идентичных клонов была бы крайне мала, что противоречит действительности. Благодаря редкости рекомбинаций нарушение любой комбинации генов будет происходить с крайне низкой частотой; в таком случае доминирующую роль в определении степени генетической изменчивости должны играть процессы клонального и «периодического» отбора. «Периодическим» отбором здесь называется явление, когда в популяции время от времени появляется новый мутантный клон с высокой приспособленностью, который быстро распространяется по популяции, вытесняя клон, существовавший в ней ранее.

В связи с этим уместно рассмотреть анализ Маруямы и Кимуры [342] эффективного размера популяции в условиях локальной элиминации и реколонизации (разд. 3.3). Авторы показали, что если вид (популяция) состоит из n гаплоидных воспроизводящихся бесполом путем линий (субпопуляций), каждая из которых подвергается элиминации и последующему замещению, средняя виртуальная гетерозиготность в состоянии равновесия равна $\bar{H}_e = 2N_e v / (1 + 2N_e v)$, где N_e задается формулой

(3.9). Используя эту модель, можно показать, что, если локальная элиминация и реколонизация происходят достаточно часто, это приводит не только к существенному уменьшению эффективного размера популяций данного вида, но и в значительной мере предупреждает дивергенцию субпопуляций. В качестве примера рассмотрим гаплоидный вид, состоящий из 100 000 линий ($n=10^5$), каждая из которых основана одной особью. Хотя линия может состоять из огромного числа особей, она элиминируется в среднем один раз в 1000 поколений ($\lambda=10^{-3}$), поэтому ее среднегармоническая численность составляет всего лишь $\bar{N}=100$. Предположим, что скорость мутирования на поколение $\nu=10^{-8}$; тогда $N_e \approx 5 \cdot 10^7$, а $\bar{N}_e \approx 0,5$. Это означает, что среднее значение виртуальной гетерозиготности этого вида составляет лишь 50% от такового у *E. coli*. В то же время в любой момент времени этот вид может состоять из огромного числа (скажем, 10^{20}) особей. Можно также показать, что в этом случае генетическая изменчивость почти полностью обусловлена различиями между линиями.

Из наблюдений Силандера и Левина [507], а также математического анализа Маруямы и Кимуры [342] четко следует, что генетическую структуру популяций *E. coli* следует рассматривать как набор размножающихся бесполом путем линий. Эти линии часто подвергаются случайной выборке вследствие периодического отбора. Мы можем считать, что каждая из этих линий возникла от одной особи (бактерии) в результате появления мутации, обеспечивающей повышенную конкурентоспособность этой особи, и ее потомки быстро размножались бесполом путем, дав начало бесчисленному числу особей. Такая линия может в конечном итоге вымереть, заменившись новой «периодической» линией. Кишечник млекопитающих можно рассматривать как своего рода «хемостат», в котором непрерывно происходит периодический отбор. Вероятно, в действительности эффективный размер популяций *E. coli* не так уж велик, и тогда критика Милкмэном теории нейтральности ни на чем не основана. Замечательную проницательность в этом отношении проявил Ней [402], предположив, что эффективный размер популяции *E. coli* — если рассматривать ее за долгий период эволюционной истории этого вида — должен быть гораздо меньше 10^{10} , что противоречит утверждению Милкмэна о том, что размер популяции *E. coli* превышал 10^{10} по крайней мере в течение $4 \cdot 10^{10}$ поколений. «Это объясняется тем, что в одних условиях колонии *E. coli* быстро растут, а в других быстро вымирают» [402].

Недавно, используя хемостат, Дикхьюзен и Хартл [91] поставили изящные эксперименты для проверки селективной нейтральности аллозимов 6-ФГД (6-фосфоглюконатдегидрогеназы), кодируемых аллелями локуса *gnd* *E. coli*. Эти авторы путем

трансдукции перенесли четыре встречающихся в природе аллеля данного локуса (W^+ , $F2$, $S4$ и $S8$) на генетический фон линии *E. coli* K12 и с большой тщательностью выделили изогенные линии, несущие эти аллели. Их эксперименты по конкуренции линий в хемостате позволяли обнаружить давление отбора, если коэффициент отбора составлял не менее 0,5% в час, или около 1% на поколение (этот предел определялся ошибкой выборки). Количество глюконата в среде лимитировалось исходя из того, что глюконат в клетке фосфорилируется и становится субстратом 6-ФГД, поэтому ограничение количества глюконата в среде может привести к увеличению селективных различий между аллозимами.

Оказалось, что в пределах разрешающей способности эксперимента (т. е. 0,5% в час) данные аллели являются селективно нейтральными на генетическом фоне нормальной *E. coli* K12. Никаких свидетельств частотно-зависимого отбора также не было получено. Эти эксперименты проводились на генетическом фоне *edd*⁺. Локус *edd* кодирует фосфоглюконатдегидратазу, обеспечивая, как известно, альтернативный метаболический путь для 6-фосфоглюконата, минуя 6-ФГД.

Считалось, что небольшие различия в функциональной эффективности аллозимов могут компенсироваться путем «переключения» метаболизма субстрата на другой путь, где скорость его превращения немного больше. Исходя из этого, были проведены эксперименты по отбору на генетическом фоне *edd*⁻ (где альтернативный метаболический путь устраняется), и отбор аллозимов был обнаружен. Более того, отбор проходил именно в том направлении, который ожидался исходя из измеренной величины K_M для данных аллозимов (отметим, что чем меньше константа Михаэлиса K_M , тем выше сродство фермента к субстрату). Как считали Дикхьюзен и Хартл [91], эти результаты свидетельствуют о том, что аллозимы 6-ФГД в природных популяциях селективно нейтральны или почти нейтральны, однако всегда существует потенциальная возможность для отбора по ним. Такие возможности могут быть реализованы в соответствующих условиях среды или на соответствующем генетическом фоне, и тогда данные аллозимы могут стать исходным материалом для адаптивной эволюции. Эти авторы привели также результаты статистического теста, выполненного в их лаборатории с использованием данных Силандера и Левина по 109 природным линиям *E. coli*. Было исследовано соответствие распределения частот 15 аллелей локуса *gnd* формуле Юэнса для случая селективной нейтральности (см. гл. 8, где обсуждается теория выборок Юэнса). Оказалось, что наблюдаемое распределение не противоречит гипотезе селективной нейтральности. Хартл и Дикхьюзен [174] исследовали с использованием хемостата еще 2 аллеля *gnd* и пришли к такому же выводу.

С точки зрения теории нейтральности открытие Дикхьюзена и Хартла, состоящее в том, что для нейтральных или почти нейтральных аллелей могут существовать потенциальные возможности для отбора, реализующиеся в соответствующих условиях, имеет далеко идущие последствия. Мне хотелось бы назвать это явление «эффектом Дикхьюзена — Хартла» (мы еще вернемся к нему в следующей главе). Если этот эффект, как я и предполагаю, носит общий характер, то даже придерживающиеся крайних взглядов сторонники селекционистской теории должны признать, что аллели, кодирующие аллозимы, могут быть селективно нейтральными или почти нейтральными, не являясь в то же время эволюционно «бессмысленными».

Согласно Уиллзу ([582], с. 210), отбор по локусу октанола-дегидрогеназы у *D. pseudoobscura*, не выявляемый в нормальных полиморфных аутбредных популяциях, может быть обнаружен путем одновременного использования инбредных линий и «стрессовой среды» («stress media») с октанолом и этанолом. Если считать, что эффект Дикхьюзена — Хартла существует, то это утверждение уже не будет противоречить теории нейтральности.

В качестве еще одного примера, иллюстрирующего тот факт, что уровни генетической изменчивости гаплоидных и диплоидных организмов различаются не очень сильно, мне хотелось бы привести недавнюю работу Ямадзаки [608] по изменчивости гаплоидного печеночного мха *Conocerphalum conicum*. Ямадзаки исследовал методом электрофореза в крахмальном геле 11 локусов этого растения из 7 различных популяций в окрестности Фукуока (юг Японии). 7 из 11 изученных локусов оказались полиморфными; средняя виртуальная гетерозиготность составляла 0,167. Сходные результаты были получены ранее при изучении другого гаплоидного организма, *Neurospora intermedia* [526]. Ямадзаки, однако, отмечает, что нейроспора зачастую существует в природных условиях в состоянии гетерокариона, близком к диплоидному, поэтому *Conocerphalum* в большей мере соответствует «идеальному» гаплоидному организму. Таким образом, исследования, проведенные на *Conocerphalum*, с большей определенностью исключают сверхдоминирование как общий механизм поддержания генетической изменчивости на молекулярном уровне.

9.3. Статистические тесты селективной нейтральности

Теория нейтральности основывается на математической теории популяционной генетики, и, как указывалось в гл. 8, в ее рамках разработано множество математических моделей. Мы лучше понимаем поведение в конечной популяции именно нейтральных аллелей, чем любых аллелей, на которые действует

отбор. Отсюда следует, что теорию нейтральности можно легко проверить, проанализировав характер распределения частот аллелей в природных популяциях. Говоря языком статистики, мы считаем теорию нейтральности нулевой гипотезой и проверяем, насколько она соответствует наблюдениям. С этой целью предлагалось и было использовано несколько тестов, однако выяснилось, что каждый из них имеет свои недостатки. Тем не менее до сих пор применение таких тестов не дало никаких четких доказательств неправильности теории нейтральности. Напротив, тест Нея свидетельствовал в пользу этой теории, о чем я и расскажу чуть позже. Однако, прежде чем переходить к рассмотрению теста Нея, необходимо подчеркнуть, что существование вредных аллелей никоим образом не противоречит теории нейтральности. Как отмечалось в гл. 8, возможно, что вредные аллели возникают в 10 раз чаще, чем селективно нейтральные, хотя встречаемость вредных аллелей в популяциях ограничена интервалом низких частот (например, менее 10%). Это означает, что простое тестирование на строгую нейтральность и соответственно на отсутствие отбора не может привести к разрешению противоречий между «нейтрализмом» и «селекционизмом».

Чтобы лучше уяснить суть дела, воспользуемся обозначениями Ли [311]. Он различает две гипотезы: гипотезу «полной нейтральности» (pan-neutrality), обозначаемую H_p , согласно которой все аллели в популяции селективно нейтральны, и гипотезу «нейтральных мутаций» (neutral mutation, H_n), которая утверждает, что генетическая изменчивость, или гетерозиготность популяций, в основном обусловлена нейтральными или почти нейтральными мутациями. Само собой разумеется, что я придерживаюсь теории H_n , а не H_p .

Однолокусный тест Юэнса и Уоттерсона

В гл. 8 я уже рассматривал теорию выборок Юэнса [97] и упоминал о его предложении проверить селективную нейтральность, используя приближенный статистический метод [см. уравнение (8.31) и последующее обсуждение]. Работая над решением проблемы в том же направлении, Уоттерсон [572, 574, 575] построил статистическую теорию для более точной проверки селективной нейтральности аллелей с использованием выборочной гомозиготности

$$\hat{F} = \sum_{i=1}^k n_i^2 / n^2, \quad (9.1)$$

где n — размер выборки (т. е. общее число генов), k — число разных аллелей в выборке, а $n_1, n_2, \dots, n_k$ — число представителей этих аллелей в выборке ($n = n_1 + n_2 + \dots + n_k$). Как и

Юэнс, Уоттерсон использует число аллелей (k) для оценки неизвестного параметра $\theta = 4N_e v$, а затем с помощью этого параметра оценивает $\hat{F}$. Один из примеров, приводимых Уоттерсоном [572], — выборка размером $n=582$ для локуса эстеразы-2 *D. willistoni*. В этой выборке представлены 7 аллелей, причем $n_1=559$, $n_2=11$, $n_3=7$, $n_4=2$, $n_5=n_6=n_7=1$. Используя свой метод, Уоттерсон показал, что оценка $\hat{F}$ значительно превышает ожидаемую в условиях селективной нейтральности величину и что вероятность получения такого большого значения F в рамках теории нейтральности равна только 0,0069. Отсюда он делает вывод, что «нейтральность в данном случае отсутствует, а отбор, вероятно, действует в пользу гомозигот». Это означает, что сверждоминирования здесь нет, и предполагает существование слабо вредных аллелей. Аналогичный тест в случае *D. tropicalis* не обнаружил достоверных отклонений от гипотезы нейтральности.

Хотя тест Уоттерсона (и Юэнса) хорошо обоснован математически, при его применении к реальным данным возникают некоторые затруднения. Он принадлежит к типу тестов, рассматривающих каждый локус (а также каждую популяцию) по отдельности. Ли [311] отмечает, что подобный «однолокусный» подход обычно имеет три недостатка: 1) большую среднеквадратичную ошибку параметра $\theta (=4N_e v)$; 2) малую статистическую мощность; 3) невозможность использовать данные по мономорфным локусам.

Кроме того, к определенным трудностям в использовании тестов Юэнса и Уоттерсона приводят существование слабо вредных аллелей и невозможность выявления абсолютно всех различий между аллелями с помощью электрофореза, поскольку эти тесты основаны на предположениях о селективной нейтральности всех аллелей (гипотеза полной нейтральности H_p) и о том, что каждая мутация приводит к возникновению нового аллеля, отличающегося от аллелей, существовавших в популяции ранее (модель бесконечного числа аллелей). Ли [311] показал на числовых примерах, что, если средняя гетерозиготность популяции обусловлена в основном селективно нейтральными аллелями, но в популяции содержится много вредных аллелей, каждый из которых имеет низкую частоту, то метод Юэнса, основанный на оценке θ по нескольким различным аллелям, непригоден для проверки гипотезы «нейтральных мутаций» H_n , хотя он может быть пригоден для проверки гипотезы «полной нейтральности» H_p .

Эти трудности можно преодолеть, объединив данные по всем имеющимся локусам и используя для оценки θ среднюю гетерозиготность $\bar{H}_e$ [401, 439, 440]. Это дает ряд преимуществ по сравнению с подходом Юэнса — Уоттерсона при проверке пра-

вильности гипотезы нейтральности H_n (в которой допускается существование неблагоприятных аллелей), особенно если скорость мутирования нейтральных аллелей в разных локусах считается неодинаковой [408].

Тест Нея

Тест Нея основан на следующих соображениях [401]: если оценивать параметр $\theta = 4N_e\mu$ исходя из средней гетерозиготности $\bar{H}_e$ по формуле (8.20) $\theta = \bar{H}_e / (1 - \bar{H}_e)$ и использовать его для расчета дисперсии гетерозиготности (σ^2_H) по формуле Стюарта (8.16), то полученное таким образом значение дисперсии можно сравнить с соответствующей наблюдаемой величиной с целью проверки теории нейтральности. Для этого теста можно оценить дисперсию гетерозиготности по разным локусам в пределах одного вида, и он относится к упоминавшемуся классу методов, в основе которых лежит объединение данных по большому числу локусов. Данный тест широко применялся Неем и сотр. [122, 410], при этом были получены весьма убедительные свидетельства в пользу теории нейтральности.

Отметим, что наблюдаемая величина дисперсии, вычисленная по данным о нескольких локусах, включает в себя выборочную дисперсию генных частот, которую необходимо вычесть (см. [418]). С другой стороны, формула Стюарта для теоретического значения дисперсии (σ^2_H) не учитывает различие в скоростях мутирования у различных локусов, поэтому при вычислении ожидаемой дисперсии желательно учитывать изменчивость скорости мутирования.

Тест Нея был использован для анализа данных по 129 видам и 6 различающимся подвидам животных, для каждого из которых рассматривалось не менее 20 локусов; выборки включали в себя по крайней мере 30 особей.

На рис. 9.4 представлены результаты этого исследования, касающиеся данных по группе позвоночных, состоявшей из 95 видов и одного подвида. Наблюдаемые зависимости между средней гетерозиготностью и межлокусной дисперсией обозначены следующим образом: квадратики — млекопитающие, кружки — пресмыкающиеся, треугольники — рыбы, крестики — земноводные. Соответствующая зависимость, полученная из формулы Стюарта для дисперсии и формулы Кимуры — Кроу (8.15) для средней гетерозиготности, представлена сплошной линией. Видно, что экспериментальные и теоретические данные удовлетворительно соответствуют друг другу. Штриховая кривая на этом графике отвечает теоретической зависимости, полученной в рамках модели ступенчатого мутирования (разд. 8.5), а штрихпунктирная — теоретической зависимости в рамках модели бесконеч-

ного числа аллелей, в которую включено предположение о варьировании скорости мутирования (с коэффициентом вариации $\alpha=1$) [см. формулу (8.22) и последующие объяснения]. Отметим, что большинство экспериментальных точек не выходит за пределы 95%-ного доверительного интервала, ограниченного пунктирными кривыми. Аналогичным образом авторы проанали-

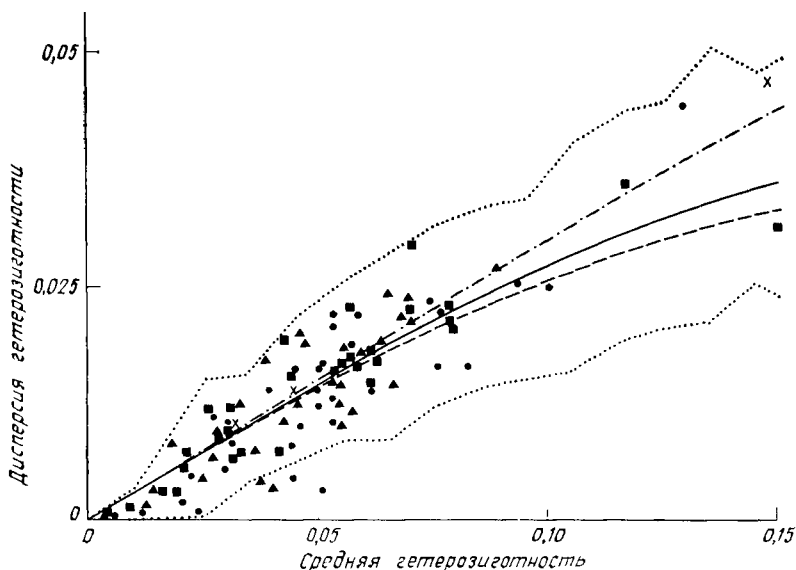


Рис. 9.4. Зависимость между средней гетерозиготностью и межлокусной дисперсией для разных видов позвоночных. — теоретическая зависимость, полученная с помощью модели бесконечного числа аллелей; --- теоретическая зависимость для модели ступенчатого мутирования; —. — теоретическая зависимость для модели бесконечного числа аллелей с переменной скоростью мутирования (коэффициент вариации скорости мутирования=1,0); ...95%-ный доверительный интервал дисперсии; ■ — млекопитающие; ● — пресмыкающиеся; ▲ — рыбы; × — земноводные. (По [122].)

зировали данные по 34 видам и 5 подвидам беспозвоночных. Здесь соответствие между наблюдаемыми и теоретическими данными также было хорошим, однако наблюдаемые значения обнаружили тенденцию к превышению теоретических значений, полученных в предположении о постоянной скорости мутирования. Это объясняется более высокой средней гетерозиготностью беспозвоночных по сравнению с позвоночными, в связи с чем влияние межлокусных различий в темпах мутирования, приводящее к увеличению наблюдаемой величины дисперсии, выражено более ярко. Действительно, теоретические значения, полученные из модели, учитывающей изменчивость скорости мутирования, пре-

красно соответствовали экспериментальным данным — гораздо лучше, чем в предыдущем случае.

Сходным образом производится оценка дисперсии и гетерозиготности по гомологичному локусу у различных видов. В этом случае скорость мутирования ν должна быть одинаковой для разных видов, однако эффективный размер популяции N_e может изменяться от вида к виду. Кроме того, в том случае, когда после дивергенции родственных видов прошло мало времени, между гетерозиготностями этих видов может существовать корреляция, что также уменьшает межвидовую дисперсию. Учитывая эти трудности, Фуэрст и др. [122] проанализировали данные по 125 локусам для различных позвоночных и получили очень хорошее соответствие с теоретическими предсказаниями. Что касается беспозвоночных (дрозофила), то здесь имелась соответствующая информация по 47 локусам. В этом случае экспериментальные результаты в основном хорошо согласовались с теоретическими, однако для некоторых групп видов *Drosophila* между ними выявилось достоверное расхождение, причем наблюдаемое значение дисперсии было по неизвестной причине ниже ожидаемого.

Чтобы проверить адекватность теории нейтральности, Ней с сотр. [44] исследовали закономерности генетической дифференциации между популяциями. Эти авторы проанализировали зависимость между генетической идентичностью I и коэффициентом корреляции r гетерозиготности различных популяций или видов, с тем чтобы проверить, соответствует ли ее наблюдаемая взаимосвязь ожидаемой на основе теории нейтральности. Как было показано в предыдущей главе [см. формулу (8.116)], параметр I служит мерой генетического расстояния между популяциями; в условиях полной изоляции он уменьшается со скоростью 2ν на поколение. Поскольку скорость мутационного процесса может быть неодинаковой для разных локусов, в работе [44] величину I вычисляли по формуле, учитывающей межлокусную изменчивость. Коэффициент корреляции r служит мерой корреляции между двумя популяциями по гетерозиготности изученных локусов. В условиях полной изоляции r уменьшается со скоростью $4\nu + 1/N_e$ на поколение [313]:

$$r = \exp[-(4\nu + 1/N_e)t], \quad (9.2)$$

где t — время в поколениях, прошедшее после дивергенции.

Наблюдаемая зависимость между I и r была проанализирована для 4 групп животных: млекопитающих, костистых рыб, пресмыкающихся и *Drosophila*. При определении теоретической зависимости, основанном на теории нейтральности, для позвоночных было принято, что $\bar{M} = 0,06$ и $\bar{H} = 0,05$, тогда как для *Drosophila* были взяты значения $\bar{M} = 0,28$ и $\bar{H}_e = 0,17$ ($\bar{M}$ — среднее $4N_e\nu$, $\bar{H}_e$ — средняя гетерозиготность на локус). Соответствие

между наблюдаемыми и ожидаемыми величинами для млекопитающих и костистых рыб было хорошим. Для пресмыкающихся зависимость между I и r однозначно установить не удалось, хотя соответствие наблюдаемых величин теоретическим во многих случаях оказалось удовлетворительным. С другой стороны, в случае *Drosophila* наблюдаемые значения коэффициента корреляции были явно выше, чем ожидалось исходя из теории нейтральности. Авторы не смогли дать четкого объяснения этому расхождению. В работе [44] был также проведен анализ зависимости между средней и дисперсией генетического расстояния. В этом случае для всех четырех групп организмов ожидаемые и наблюдаемые величины удивительно хорошо соответствовали друг другу.

Тест Левонтина — Кракауэра

Этот тест [305] основан на той идее, что, хотя отбор может действовать на разные локусы неодинаковым образом, влияние инбридинга должно быть для всех локусов одним и тем же. Левонтин и Кракауэр предложили следующий метод проверки полиморфных локусов на нейтральность. Пусть имеется локус, представленный двумя аллелями — A_1 и A_2 . Определим коэффициент инбридинга F , используя для этого среднее ($\bar{p}$) и дисперсию (σ_p^2) частоты A_1 в популяциях данного вида, по формуле

$$F = \sigma_p^2 / [\bar{p}(1 - \bar{p})]. \quad (9.3)$$

Эта мера инбридинга соответствует величине F_{ST} Райта [594]. Определим таким образом F для всех исследуемых локусов, а затем вычислим среднее значение $\bar{F}$ и дисперсию s_F^2 полученных для разных локусов значений F . Предположим, что F определялось в n популяциях. Левонтин и Кракауэр утверждают, что в случае селективной нейтральности консервативная оценка s_F^2 равна $2\bar{F}^2/(n-1)$, и это можно использовать для проверки правильности теории нейтральности: если $k = (n-1)s_F^2/\bar{F}^2 > 2$, теория нейтральности несправедлива. Другими словами, если s_F^2/σ_F^2 [где теоретическая дисперсия $\sigma_F^2 = 2\bar{F}^2/(n-1)$] много больше единицы, значения F считаются слишком гетерогенными, чтобы их можно было объяснить без предположения о наличии естественного отбора.

По-видимому, первым, кто высказал мысль об использовании степени единообразия значений F по разным локусам для оценки роли генетического дрейфа, был Кавалли-Сфорца [40], который вычислил F для различных систем групп крови у множества популяций человека. Обработав данные Кавалли-Сфорца своим методом, Левонтин и Кракауэр получили значение $s_F^2/\sigma_F^2 = 10$, что, по их мнению, свидетельствовало об очень высо-

кой гетерогенности значений F для полиморфных генов популяций человека по всему миру и, следовательно, о том, что наблюдаемые частоты значительной части полиморфных генов человека обусловлены действием естественного отбора.

Робертсон [491] подверг сомнению справедливость теста Левонтина — Кракауэра, показав, что при наличии корреляции между популяциями (скажем, из-за миграции или недавнего разделения популяций), характерной, например, для человека, оценка $k=2$ может быть сильно заниженной. Робертсон отметил, что для выборки из 60 популяций, исследованной Левонтином и Кракауэром, правильное значение σ^2_F в 6 раз выше значения, полученного этими авторами. Таким образом, поскольку наблюдаемая межлокусная дисперсия в 10 раз превышает величину, следующую из расчетов Левонтина и Кракауэра, возникают некоторые сомнения по поводу правильности их утверждения о гетерогенности значений F (более детальное рассмотрение вопроса об ожидаемых значениях k в случае корреляции между популяциями дано в работе [98]). Тест Левонтина — Кракауэра был подвергнут критическому анализу Неем и Маруямой [414], которые провели методом Монте-Карло моделирование на ЭВМ генетических процессов в подразделенных популяциях и показали, что наблюдаемые значения k обычно бывают гораздо больше двух. В ответ на эту критику Левонтин и Кракауэр [306] перечислили «за» и «против» метода проверки величины F на гетерогенность. В частности, в качестве примера правильного использования их методики они привели ее применение в случае временной (а не географической) изменчивости генных частот в популяциях. В этой работе ими были использованы данные Кримбаса и Цакаса [285] по двум полиморфным локусам маслинной мухи, *Dacus oleae*. Различие временных коэффициентов инбридинга в случае этих двух локусов не было статистически значимым. Позднее тест Левонтина — Кракауэра применялся для анализа временной (в течение 3 лет) изменчивости генных частот пяти электрофоретически различимых локусов в каждой из четырех популяций полевки (*Microtus ochrogaster*) флуктуирующей численности [124]. Все наблюдаемые значения величины s^2_F/σ^2_F были близки к 1. Авторы сделали вывод, что главной причиной изменения генных частот в популяции полевки являются «неселективные факторы», в особенности дрейф генов.

Другие тесты

Кимура и Ота [256, 257] исследовали зависимость между средней гетерозиготностью ($\bar{H}_e$) и долей полиморфных локусов ($P_{\text{поли}}$), используя формулу $P_{\text{поли}} = 1 - q^{H/(1-H)}$. Эта формула была получена в рамках модели бесконечного числа аллелей, при-

чем q обычно принималось равным 0,01 или 0,05 [см. формулу (8.9) и ее обсуждение в гл. 8]. Применяв эту формулу к данным по нескольким организмам, авторы показали, что наблюдаемые значения довольно хорошо согласуются с теоретическими. Позднее этот же тест был использован для анализа гораздо более обширного материала [122], причем полученные результаты полностью соответствовали теории нейтральности.

Джонсон [196] для проверки теории нейтральности предложил исследовать зависимость между фактическим числом различных аллелей $n_{\text{ф}}$ и отношением эффективного числа аллелей к их фактическому числу, т. е. $n_e/n_{\text{ф}}$, где n_e — величина, обратная сумме квадратов аллельных частот. Он пришел к выводу, что если теория нейтральности верна, $n_e/n_{\text{ф}}$ должно уменьшаться с увеличением $n_{\text{ф}}$, поскольку равновесное распределение частот аллелей будет сильно скошено (из-за наличия при больших $n_{\text{ф}}$ большого числа аллелей с низкой частотой). Исследовав эту зависимость для трех групп данных по дрозофиле, Джонсон представил «совершенно четкие» результаты, свидетельствующие об увеличении $n_e/n_{\text{ф}}$ с увеличением $n_{\text{ф}}$ во всех трех случаях в противоположность предсказанию теории нейтральности. Ямадзаки и Маруяма [610] провели повторные исследования, используя гораздо более обширные данные и точные значения теоретических величин, основанные на модели бесконечного числа нейтральных аллелей. В отличие от Джонсона они выявили прекрасное соответствие наблюдаемых данных предсказаниям теории нейтральности. Ямадзаки и Маруяма отметили также, что Джонсон исключал из рассмотрения аллели, имеющие частоту менее 0,01, и это могло привести его к неправильным выводам.

Ямадзаки и Маруяма [609, 611] разработали свой подход к проверке селективной нейтральности полиморфизма белков. Он основан на наблюдении Маруямы [334], состоящем в том, что ожидаемая сумма гетерозигот, образованных мутацией с момента ее появления до момента фиксации или утраты, «инвариантна», т. е. не зависит от структуры популяции. В этом подходе используется также следующая теоретическая закономерность: при условии возникновения необратимых мутаций и в рамках модели бесконечного числа сайтов стационарное распределение частот нейтральных мутаций задается формулой $\Phi(x) = 4N_e v/x$, которая является частным случаем формулы (8.97), когда $G(x) = 1$ и $v_{\text{гам}} = v$, где v — скорость мутирования на локус, а x — частота мутантного аллеля. Умножая $\Phi(x)$ на $2x(1-x)$, получаем выражение $8N_e v(1-x)$, определяющее относительные частоты гетерозигот. Ямадзаки и Маруяма рассмотрели также случай, когда в популяции имеет место отбор. Они показали, что, если мутация является безусловно благоприятной, соответствующая частота гетерозигот примерно постоянна, тогда как в случае

явно неблагоприятной мутации распределение имеет минимум при $x=0,5$ и достигает максимума при $x=0$ или $x=1$. В случае сверхдоминантной мутации построить точное распределение довольно трудно, однако авторы для простоты предположили, что все равновесные частоты равновероятны, т. е. распределены равномерно (на самом деле аллели с равновесной частотой, близкой к 0,5, сохраняются чаще). В этом случае распределение гетерозиготности становится колоколообразным с максимумом при $x=0,5$. Одна из трудностей такого подхода заключается в следующем: когда мы рассматриваем гетерозиготу (скажем, AA'), мы не знаем, какой именно из аллелей — A или A' — является мутантным. Чтобы обойти эту трудность, Ямадзаки и Маруяма предложили объединить обе половины шкалы генных частот, ограничившись интервалом от 0 до 0,5; при этом значения ординат для точек x и $1-x$ суммируются. Распределение гетерозиготности нейтральных мутаций в этом интервале частот становится равномерным, как и в случае благоприятных мутаций.

Затем Ямадзаки и Маруяма построили график зависимости суммарной наблюдаемой гетерозиготности (ордината) от частоты гена (абсцисса). Другими словами, они рассмотрели относительный вклад аллелей каждого класса частот в полную гетерозиготность. При этом они исключали из каждого локуса один случайно выбранный аллель, чтобы внести поправку на отклонения, обусловленные тем, что частоты аллелей одного локуса не являются независимыми. Используя все имеющиеся данные (442 аллеля по 145 белкам 16 видов; см. работу [609]), Ямадзаки и Маруяма получили результаты, хорошо согласующиеся с ожидаемыми на основе равномерного распределения гетерозиготности, что возможно как в случае нейтральности, так и в случае явного благоприятного эффекта мутации. Из этих двух возможностей авторы отдают предпочтение гипотезе селективной нейтральности, поскольку маловероятно, что большинство новых мутаций обладает большей приспособленностью, чем существовавшие ранее аллели.

Метод Ямадзаки и Маруямы был подвергнут критическому анализу Юэнсом и Фельдманом [100] на том основании, что при больших $4N_e\mu$ используемая авторами формула становится неточной. Однако это замечание несущественно, поскольку для нейтральных аллелей величину $4N_e\mu$ можно считать намного меньшей единицы. Позже Ямадзаки и Маруяма повторили свой анализ, используя более обширный материал (1530 аллелей 600 полиморфных локусов [612]), однако распределение гетерозиготностей оказалось таким же, снова соответствуя теории нейтральности.

Айала и Гилпин [9] подвергли метод Ямадзаки и Маруямы критике, утверждая, что найденное ими теоретическое распределение относительного уровня гетерозиготности в предполо-

жении о наличии сверхдоминирования неверно, и потому при равенстве аллельных частот наблюдаемое распределение гетерозиготности хорошо согласуется с моделью сверхдоминирования. Однако, отвечая на критические замечания Айалы и Гилпина, укажем, что некорректно, как это делают Айала и Гилпин, произвольно полагать, что все аллельные комбинации для данного числа аллелей равновероятны, т. е. считать, что все точки на гиперплоскости, определенной как $Y_1 + Y_2 + \dots + Y_n = 1$ (где Y_i — частота i -го аллеля), имеют равную вероятность.

Кроме того, Айала и Гилпин [10] высказали следующие критические замечания по поводу теории нейтральности: если эта теория верна, распределение генетического расстояния D между таксонами должно быть близким к нормальному или колоколообразному (причем вероятность того, что $D=0$, должна быть исчезающе мала), тогда как в действительности наблюдаемое распределение генетических расстояний между таксонами в группе *D. willistoni* во всех случаях является U-образным, что противоречит теории нейтральности (см. также [7]). В своих теоретических расчетах эти авторы численно решали диффузионное уравнение Фоккера — Планка и находили вероятностное распределение расстояния D при $t=0,1N, 0,2N$ и т. д. между двумя изолированными популяциями размером N каждая, предполагая, что они дивергировали от одной и той же популяции в момент $t=0$. Здесь $D = [\Sigma (X_{1i} - X_{2i})^2 / 2]^{1/2}$, где X_{1i} и X_{2i} — частоты i -го аллеля в первой и второй популяциях, а суммирование производится по всем аллелям данного локуса.

Однако, как отмечали Ней и Татено [419], расчеты Айалы и Гилпина [10] основаны на двух предположениях, не соответствующих действительности. Во-первых, эти авторы считают, что в момент дивергенции генные частоты одинаковы для всех локусов. В действительности же они сильно варьируют от локуса к локусу: некоторые локусы мономорфны, тогда как другие в разной степени полиморфны. Во-вторых, Айала и Гилпин предполагают отсутствие мутирования, что тоже нереально. В их работе анализируется следующая ситуация: популяция содержит, к примеру, три аллеля с одинаковыми начальными частотами ($1/3, 1/3, 1/3$); затем рассматривается процесс, в результате которого один из этих аллелей со временем закрепляется в популяции. Мне кажется, что Айала и Гилпин не учли возможность того, что мономорфный локус остается в таком состоянии после дивергенции (т. е. $D=0$) вплоть до возникновения новых, благоприятных мутаций и распространения их до ощутимых частот в любой из двух популяций. Такой процесс может быть весьма продолжительным. Далее, даже в случае полиморфных локусов по каждому из них обычно существует наиболее распространенный аллель; для его замены другими аллелями потребуется долгое время.

Ней и Татено [419] при моделировании популяционных процессов использовали вполне реалистичное распределение генных частот разных локусов в момент дивергенции популяций и исследовали изменение величины D (определенной так, как это было сделано Айалой и Гилпином) после дивергенции под влиянием мутационного процесса и случайного дрейфа генов, положив $M=4N_e\mu=0,1$. Выяснилось, что распределения полученных таким образом значений D через $0,5N_e$ и N_e поколений являются L-образными с максимумом в окрестности $D=0$, т. е. резко отличаются от колоколообразных теоретических распределений Айалы и Гилпина.

Ней и Татено проанализировали также процесс изменения коэффициента генетической идентичности I [см. формулу (8.116)] в рамках теории нейтральности и получили распределение, хорошо соответствующее наблюдаемым распределениям межрасовых и межвидовых генных различий по ферментным локусам у *Drosophila*. Эти результаты показывают, что в противоположность утверждению Айалы и Гилпина [10], распределение генных замещений в данных локусах можно объяснить на основе теории нейтральности.

9.4. О некоторых альтернативных гипотезах

Как показано в двух предыдущих разделах, теория нейтральности белкового полиморфизма с честью выдержала различного рода экспериментальные и статистические проверки. Если прибавить к этому успешное объяснение ею множества экспериментальных фактов, касающихся эволюционных изменений на молекулярном уровне, то станет ясно, что теория нейтральности устоит и в дальнейшем, хотя, несомненно, будет видоизменяться и совершенствоваться. Здесь мне хотелось бы обсудить некоторые альтернативные гипотезы поддержания генетической изменчивости на молекулярном уровне.

Гипотеза сверхдоминирования

Я уже привел достаточно доказательств, свидетельствующих о том, что сверхдоминирование не может быть главной причиной полиморфизма белков. Роковым ударом по представлениям о сверхдоминировании как универсальной причине поддержания полиморфизма белков явилось открытие, состоящее в том, что уровень генной изменчивости (виртуальной гетерозиготности) у гаплоидных организмов так же высок, как и у диплоидных. В этом разделе я представляю дополнительные данные, свидетельствующие против гипотезы сверхдоминирования. Используя численные и аналитические методы, Левонтин и др. [303] показали, что превосходство гетерозигот по приспособ-

ленности не может быть единственной причиной поддержания множества аллелей данного локуса (они рассматривали только детерминистические процессы). Согласно этим авторам, даже в случае, когда все гетерозиготы обладают более высокой по сравнению с гомозиготами приспособленностью, доля наборов приспособленностей, приводящих к стабильному полиморфизму более чем 6 или 7 аллелей, является исчезающе малой: после того как число аллелей превысит три, условия существования стабильного полиморфизма быстро становятся очень жесткими. Даже когда эти условия выполняются, распределение частот аллелей полиморфного локуса при равновесии становится почти равномерным. С другой стороны, наблюдаемые распределения аллелей полиморфного локуса в случае большого числа аллелей (локус *Xdh* у *Drosophila*; см. [65, 516]) в высшей степени асимметричны.

В настоящее время накопилось много данных, противоречащих гипотезе о сверхдоминировании как о механизме поддержания полиморфизма белков, и их необходимо учитывать в популяционно-генетических исследованиях.

Теория Фрэнклина — Левонтина о многолокусном неравновесии по сцеплению

В статье под названием «Является ли ген единицей отбора?» Фрэнклин и Левонтин [121] рассмотрели модель, согласно которой в хромосоме плотно «упаковано» большое число сверхдоминантных локусов с мультипликативным влиянием на приспособленность. Авторы выявили следующее замечательное свойство данной модели: в популяции возникает такое сильное неравновесие по сцеплению (linkage disequilibrium), что в ней имеются в основном только два взаимодополняющих типа хромосом. Например, в одной серии модельных экспериментов авторы использовали 36 сверхдоминантных локусов в хромосоме с частотой рекомбинаций между прилежащими локусами, равной 0,0025, и 10%-ным уменьшением приспособленности в гомозиготном состоянии по отдельному локусу. Эффективный размер популяции предполагался равным 400. Исходные значения генных частот были одинаковыми (т. е. 0,5) для всех локусов, а комбинирование генов между локусами происходило случайным образом. В одной из повторностей — в 300-м поколении — популяция состояла в основном из двух типов хромосом, обозначенных

011 010 110 ... 101

и

100 101 001 ... 010

которые имели примерно одинаковые частоты. Эти два типа хромосом составляли в популяции около 83%, а на долю всех остальных типов, вместе взятых, приходилось только 17%.

В таком состоянии равновесия средняя приспособленность популяции ($\bar{w}$) становится равной 0,4, что значительно превышает величину $\bar{w}=0,16$, ожидаемую при случайном комбинировании генов. Если частота рекомбинации между смежными локусами r была равна 0,005 или выше, такого ярко выраженного эффекта «поляризации» не наблюдалось. Таким образом, для его появления сцепление должно быть достаточно сильным. Впрочем, при полном сцеплении ($r=0,0$) заметного повышения приспособленности тоже не происходило.

Весьма примечательно, что $r=0,0025$ — это вполне реальное расстояние на генетической карте между двумя соседними расщепляющими локусами у *Drosophila*. Фрэнклин и Левонтин отметили еще одно очень интересное свойство этой модели: при увеличении в хромосоме числа расщепляющихся локусов, но постоянной длине карты и эффекте инбридинга (т. е. уменьшении приспособленности аллеля, когда он находится в гомозиготном состоянии) средняя корреляция между генами в хромосоме существенно не изменялась (математический анализ см. в [517]). Другими словами, это свойство системы не зависит от числа генов; главную роль играет хромосома в целом, а не отдельные локусы. Если в распределении участвует очень большое число локусов, влияние на приспособленность гомозиготности по одному локусу очень мало, и это свойство модели делает ее более адекватной. С другой стороны, в экстремальном случае аллели в этой модели образуют только две группы, и в этом смысле модель напоминает систему сбалансированных леталей. Если подобный эффект «поляризации» действительно существует в природных популяциях, концепция «тяжелого сегрегационного груза» при наличии большого числа сверхдоминантных локусов является несостоятельной; этот груз в условиях случайного скрещивания не намного превышает 0,5.

Фрэнклин и Левонтин [121] проанализировали также другие ситуации: асимметричную приспособленность двух гомозигот по каждому локусу; локусы, неравномерно распределенные по хромосоме; отсекающий отбор. Они убедились, что выявленная ими закономерность справедлива и в этих случаях. Отсюда ими был сделан вывод, что теория популяционной генетики, считающая хромосому единым целым, больше соответствует действительности, чем традиционная теория, ограничивающаяся рассмотрением генных частот по отдельным локусам. Это утверждение было вновь приведено Левонтином [302] в его книге «Генетические основы эволюции», где в заключение он проводит аналогию между теорией, считающей хромосому единым целым, и хорошо известным законом термодинамики $PV=nRT$, согласно которому между давлением, температурой и объемом системы существует простая связь, хотя все эти параметры оп-

ределяются состоянием огромного множества молекул, каждая из которых находится в хаотическом движении. Создается впечатление, что, по мнению Левонтина, этот подход открывает новую эру в популяционной генетике.

Хотя все сказанное выше представляет несомненный интерес, вопрос о существовании в природе подобного ярко выраженного неравновесия по сцеплению остается открытым. Действительно, по мере появления новых данных становится все более очевидно, что между локусами, кодирующими ферменты у *Drosophila*, не существует практически никакого неравновесия по сцеплению, за исключением случаев, когда эти локусы ассоциированы с инверсиями (структурными изменениями хромосомы, при которых участок хромосомы оказывается повернутым на 180°).

То, что неравновесие по сцеплению в случае изозимных локусов редко встречается в природных популяциях *Drosophila*, было показано в обширных исследованиях Мукаи и др. [375, 378, 379], Лэнгли и др. [292] и Лукаса и др. [315, 316]. Неравновесие по сцеплению не было обнаружено даже между такими тесно сцепленными локусами, как *Odh* (октанолдегидрогеназа) и *Est-C* (эстераза-С), частота рекомбинации между которыми составляет 0,0058 [378]. Таким образом, хотя теория Фрэнклина — Левонтина может быть применена в некоторых случаях для объяснения неравновесия по сцеплению на очень коротком участке хромосомы, в качестве общей теории поддержания полиморфизма белков она неприемлема. Это хороший пример того, как прекрасная научная гипотеза не выстояла под натиском неумолимых фактов.

Между прочим, несомненное неравновесие по сцеплению было выявлено в некоторых случаях между генами, кодирующими изозимы, и полиморфными инверсиями. Эти факты, однако, наиболее успешно объясняются эффектом основателя, т. е. историческими причинами и случайным дрейфом генов. Ней и Ли [413] провели математическое исследование проблемы неслучайной ассоциации аллелей, кодирующих изозимы, и инвертированных хромосом в популяциях конечного размера. Используя метод диффузионных уравнений, эти авторы оценили возможный уровень неслучайной ассоциации, т. е. разность частот данного аллеля в инвертированных и неинвертированных хромосомах. Согласно проведенному ими анализу, достоверная неслучайная ассоциация имеет место при малых (k примеру, меньше 1) $4N_1r$, где N_1 — эффективное число инвертированных хромосом, а r — скорость, с которой аллель, находящийся в инвертированной хромосоме, заменяется на аллель, находящийся в неинвертированной хромосоме. Ней и Ли показали, что стандартная ошибка неслучайной ассоциации зачастую превышает среднее.

Ранее Пракаш и Левонтин [472] отметили, что для локусов *Pt-10* и *Amylase* характерна прочная ассоциация определенных аллелей с разными типами инверсий («Стандард», «Санта Крус») третьей хромосомы *Drosophila pseudoobscura* независимо от географического происхождения линий. Кроме того, они показали, что у вида-двойника, *D. persimilis*, инверсии третьей хромосомы которого принадлежат к типу Стандард (или произошли от него), аллельный состав очень близок к аллельному составу Стандард у *D. pseudoobscura*. Учитывая, что возникновение этих типов инверсий могло произойти по крайней мере несколько миллионов поколений назад, эти авторы пришли к выводу о «коадаптированности» состава инверсий, при котором введение новых аллелей исключено, что согласуется с идеями Добржанского [83]. Однако Ней и Ли [413], проведя более тщательный математический анализ этой проблемы, показали, что наблюдения Пракаша и Левонтина [472] легко объяснимы в рамках теории нейтральности.

Экологические теории, предполагающие адаптивную стратегию

Если, как утверждают сторонники селекционистской теории, полиморфизм белков и других молекул является адаптивным, то между генетической изменчивостью и некоторыми факторами окружающей среды должна существовать корреляция. Действительно, ряд селекционистов придерживается мнения, что разнообразие генов лежит в основе адаптивной стратегии вида, направленной на освоение окружающей среды, гетерогенной во времени и пространстве.

Согласно одной популярной гипотезе (Левинс [301]), генетическая изменчивость должна быть выше в нестабильных условиях среды по сравнению со стабильными. Эта гипотеза представляется весьма правдоподобной; сходные предположения высказывались другими авторами (обзор см. в [523, 553]). Поскольку глубоководные моря являются ареалом с самыми стабильными и гомогенными условиями среды, было высказано предположение, что генетическая изменчивость обитающих там организмов очень низка. Однако оказалось, что многие организмы, живущие на дне моря на больших глубинах, характеризуются весьма высоким уровнем генетической изменчивости [15, 136]. Таким образом, это предсказание оказалось совершенно неверным. Данная гипотеза была проверена также Сомеро и Суле [522] (назвавшими ее «гипотезой изменчивости ниш»), которые рассмотрели предположение о том, что для рыб, обитающих в условиях изменяющейся температуры, характерен более высокий уровень полиморфизма белков, и опровергли его.

Несостоятельность гипотезы Левинса привела некоторых исследователей к гипотезе «стабильности трофических ресурсов» (см. [553]), согласно которой лучшая стратегия вида в условиях изобилия и стабильности пищевых ресурсов — создание аллельных вариантов, специализированных к разной пище. Именно таким образом, по мнению сторонников этой гипотезы, появляется высокая генетическая изменчивость. С другой стороны, в ситуации, когда запасы пищевых ресурсов непредсказуемо изменяются, для популяции должно быть характерно генетическое единообразие, поскольку в таких условиях лучшей стратегией является создание неспециализированного генотипа. Высказывалось также мнение [14], что у видов, обитающих в среде с изменяющимися во времени условиями, или у видов, отдельные организмы которых обитают в самых разных условиях, отбор благоприятствует функционально «гибким» аллелям, причем в экстремальных условиях для вида характерна тенденция к достижению мономорфности по этим аллелям. Такие экологические теории, предполагающие наличие разных стратегий адаптации, сами по себе кажутся мне столь «гибкими», что, привлекая какую-либо из них (или придумывая новые), можно объяснить практически любой факт, касающийся генетической изменчивости.

Возвращаясь к гипотезе Левинса, Пауэлл [470] и Мак-Доналд и Айала [320] показали, что в экспериментальных ящичных популяциях *Drosophila* уровень гетерозиготности популяции был выше в гетерогенной среде, чем в гомогенной. С тех пор эти эксперименты приводят в литературе как довод в пользу положения, что полиморфизм белков поддерживается уравновешивающим отбором. Однако Ней [406], тщательно проанализировав полученные в упомянутых работах данные, указал, что средняя скорость уменьшения гетерозиготности в этих экспериментах в обоих случаях выше, чем при наличии одного лишь генетического дрейфа. В работе Мак-Доналда и Айалы выборка была взята из природной популяции *Drosophila pseudoobscura*, имеющей гетерозиготность 0,212 (по 20 полиморфным локусам), а минимальный размер их экспериментальной популяции был не менее 500. Следовательно, максимальное уменьшение средней гетерозиготности, обусловленное случайным дрейфом генов, за 12—15 поколений (продолжительность эксперимента) должно составить $0,212 - 0,212 \cdot (1 - 1/1000)^{15} = 0,003$. Однако на самом деле средняя наблюдаемая гетерозиготность в гетерогенной среде составляла к концу эксперимента 0,195, тогда как в гомогенной среде она равнялась 0,146. Ясно, что гетерогенность среды не замедляла, а, напротив, ускорила снижение уровня гетерозиготности, что и отмечалось Неем. В аналогичном эксперименте на *Drosophila willistoni*, проведенном Пауэллом [470], уменьшение средней гетерозиготности в ге-

терогенной среде было еще более существенным. Отметим, что в обоих экспериментах уменьшение гетерозиготности в гомогенной среде было выражено сильнее. Причина этого явления неизвестна. Этот же вопрос исследовался Минавой и Берли [360] с использованием ящичных популяций *Drosophila melanogaster*. В отличие от результатов Пауэлла [470] и Мак-Доналда и Айалы [320] средняя гетерозиготность популяций, содержащихся в стабильной и изменчивой среде, достоверно не различалась, т. е. гетерогенность среды не влияла на уровень гетерозиготности. С другой стороны, в этой работе были обнаружены достоверные различия частот электроморфов в разных средах. Авторы считают, что полученные ими результаты не согласуются с теорией нейтральности. На мой взгляд, они не согласуются и с гипотезой адаптивной стратегии вида. Возможно, достоверные изменения частот электроморфов обусловлены давлением отбора на невыявленные локусы, находящиеся в неравновесии по сцеплению с изучаемыми генами. Очевидно, для окончательного решения вопроса необходимы более тщательные и детальные эксперименты. Дополнительную информацию о взаимосвязи генетического полиморфизма и гетерогенности среды читатели найдут в обзорных работах [176, 420, 523, 553].

9.5. Влияние структуры молекул и функциональных ограничений на генетическую изменчивость

С точки зрения теории нейтральности полиморфизм на молекулярном уровне представляет собой переходную фазу молекулярной эволюции. Если это действительно так, то можно ожидать, что степень полиморфизма будет сильно зависеть от структуры соответствующих молекул и налагаемых на них функциональных ограничений, что и имеет место в молекулярной эволюции (гл. 7). В связи с этим особый интерес представляет наблюдение Зоурса [619]. Он сравнил электрофоретическую изменчивость локусов, кодирующих мультимерные ферменты (т. е. ферменты, состоящие более чем из одной полипептидной субъединицы), и локусов, кодирующих мономерные ферменты, у четырех видов животных и двух видов растений. Во всех случаях уровень гетерозиготности мультимерных ферментов был ниже, чем мономерных. Было также показано, что для данного вида или популяции мультимерные ферменты обнаруживают более низкий уровень полиморфизма, чем мономерные ферменты, выполняющие сходную функцию. Данные о достоверно более высоком уровне гетерозиготности (а также полиморфизма) мономерных ферментов по сравнению с мультимерными были затем получены и другими авторами [172,

565] (см. также [277]). Как отметил Зоурос [619], эти результаты опровергают представление о том [103], что основой сверхдоминирования на молекулярном уровне является образование гетеромультимеров — ассоциатов из двух субъединиц, кодируемых разными аллелями. С другой стороны, эти наблюдения не противоречат теории нейтральности, поскольку структурные ограничения, налагаемые на правильно функционирующие молекулы, должны быть более жесткими в случае мультимеров (из-за взаимодействия субъединиц), чем мономеров, и, таким образом, при прочих равных условиях доля нейтральных (т. е. не вредных) мутаций меньше для мультимерного фермента, чем для мономерного.

Другим интересным вопросом является взаимосвязь между молекулярной массой белка и уровнем внутривидовой изменчивости. Обширный анализ данных по полиморфизму ферментов у *Drosophila* [276, 277] показал, что для 11 димерных ферментов существует достоверная корреляция между гетерозиготностью, числом аллелей и молекулярной массой субъединиц. Коэффициент корреляции r между молекулярной массой субъединицы и наблюдаемой гетерозиготностью по каждому локусу в выборке составлял 0,413, а между молекулярной массой и средней гетерозиготностью на локус — 0,753 [276]. В обоих случаях корреляция была статистически значимой. Согласно теории нейтральности или, точнее, модели бесконечного числа нейтральных аллелей, средняя гетерозиготность $\bar{H}_e = 4N_e v_0 / (1 + 4N_e v_0)$. Здесь v_0 — скорость мутирования для нейтральных аллелей, которая предполагается пропорциональной размеру гена и, следовательно, должна быть пропорциональна молекулярной массе кодируемого данным геном белка. Это означает, что средняя гетерозиготность должна сильно коррелировать с молекулярной массой. Таким образом, наблюдаемое более высокое значение коэффициента корреляции ($r=0,753$) согласуется с теорией нейтральности. Кроме того, в соответствии с теорией нейтральности наблюдаемая гетерозиготность по данному локусу определенного вида может сильно отличаться от ожидаемой величины $\bar{H}_e$ в результате влияния случайного дрейфа генов, что видно из характера распределения гетерозиготности на локус у родственных видов (см., например, рис. 9.3). Стохастические изменения, несомненно, приведут к уменьшению коэффициента корреляции, определенного отдельно для каждой выборки, по сравнению с коэффициентом корреляции, вычисленным для средней гетерозиготности нескольких видов. Это объясняет также, почему в работе [276] получено более низкое значение коэффициента корреляции ($r=0,463$).

Используя данные по 6 группам организмов — приматам, грызунам, пресмыкающимся, саламандрам, рыбам и дрозофиле, — Ней и др. [411] провели более подробный и тщательный

анализ взаимосвязи между молекулярной массой субъединиц и генетической изменчивостью белков. Были рассмотрены в основном мономерные и димерные белки, а также некоторые три- и тетрамерные. Авторы показали, что во всех группах исследованных животных коэффициент корреляции между молекулярной массой и гетерозиготностью был положительным и что при учете неполной корреляции между молекулярной массой и скоростью мутирования корреляция в основном соответствовала величине, ожидаемой из теории нейтральности. Далее ими было показано, что при высоком уровне гетерозиготности корреляция была выше, чем при низком. Это также согласуется с теорией нейтральности. Харрис и др. [172] провели аналогичный анализ, используя данные по 87 локусам, кодирующим ферменты, у человека. При сравнении средней молекулярной массы субъединиц ферментов в случае полиморфных и мономорфных локусов значимой разницы между ними обнаружено не было. Значимых различий не было выявлено и тогда, когда локусы, кодирующие мономерные и мультимерные ферменты, рассматривались по отдельности. Действительная причина такого отсутствия корреляции между молекулярной массой субъединиц и уровнем полиморфизма неизвестна. Поскольку средний уровень гетерозиготности у человека довольно низок (около 8%) и поскольку гетерозиготность по данному локусу обнаруживает высокую степень межвидовой изменчивости, обусловленную дрейфом, не исключено, что отсутствие корреляции является случайным (см., однако, [411]). С другой стороны, по данным работы [172] частота полиморфизма у мультимерных ферментов достоверно ниже, чем у мономерных, что согласуется с приведенными выше результатами Зоурса. Кроме того, обнаружен следующий примечательный факт [172]: частота полиморфизма особенно низка для мономерных ферментов, у которых встречаются «межлокусные» молекулярные гибриды. Согласно Харрису и др. [172] (см. табл. 4 в их работе), частота полиморфизма составляет 56% для 27 локусов, кодирующих мономерные ферменты, 42% для 38 локусов, кодирующих мультимерные ферменты, не образующие «межлокусных» гибридных молекул, и лишь 9% для 22 локусов, кодирующих мультимерные ферменты, которые образуют «межлокусные» гибридные молекулы. Этот результат является особенно веским доводом в пользу теории нейтральности, так как образование межлокусных гибридов налагает структурные ограничения на белковые молекулы, в результате чего уменьшается вероятность нейтральности мутационного изменения.

Тезис сторонников нейтралистской теории о том, что увеличение функциональных ограничений, налагаемых на белковую молекулу, уменьшает внутривидовую изменчивость данного белка, вероятно, имеет широкую применимость. С его по-

мощью можно по-новому подойти к интерпретации различных наблюдений и гипотез, которые когда-то выдвигались в поддержку теории селекционизма.

Гиллеспи и Кодзима [130] провели сравнение изменчивости двух групп ферментов. В одну группу входили ферменты, участвующие в энергетическом метаболизме (группа I), в другую — ферменты, имеющие широкую субстратную специфичность (группа II). Оказалось, что для ферментов группы I характерна значительно меньшая изменчивость, чем для ферментов группы II. У *Drosophila ananassae* средняя гетерозиготность на локус для ферментов группы I составляла только 9~24% от средней гетерозиготности для ферментов группы II. Сходное исследование с использованием более обширных данных по популяциям человека, мыши и дрозофилы было выполнено Гиллеспи и Лэнгли [131]. Разделение ферментов на две группы было проведено ими несколько иначе, а именно: 1) ферменты, действующие в обычных условиях только на один субстрат (эта группа в основном совпадает с группой I в работе [130]); 2) ферменты, действующие на несколько субстратов (почти то же самое, что группа II). При таком разделении четкое и устойчивое различие между средней гетерозиготностью ферментов группы I и группы II было обнаружено не только для *Drosophila*, но также для человека и мыши. Отношение средних гетерозиготностей ферментов групп I и II составляло 0,38 у человека, 0,65 у мыши и 0,17 у *Drosophila*. Было также показано, что средняя гетерозиготность ферментов группы I приблизительно постоянна и составляет у изученных организмов около 0,05. Эти результаты Гиллеспи и Лэнгли [131] расценивали как довод в пользу своего положения о том, что полиморфизм ферментов обусловлен прежде всего отбором, действующим на изменчивость функции гена под влиянием условий среды. Однако их наблюдения легко объяснить и в рамках теории нейтральности: в случае субстрат-специфичных ферментов (группа I) функциональные ограничения сильнее, чем для ферментов, у которых субстратная специфичность отсутствует (группа II); следовательно, вероятность того, что мутационное изменение (замещение аминокислоты) будет нейтральным, для ферментов группы I меньше, чем для ферментов группы II.

В ходе дальнейшего исследования этого вопроса с привлечением данных по разным видам *Drosophila* Ямадзаки [607] выявил не только более низкую степень гетерозиготности для субстрат-специфичных ферментов (в первом случае $\bar{H}_e = 0,075$, во втором — 0,200), но также и различие между этими группами по взаимосвязи между средней гетерозиготностью и ее дисперсией. Используя формулу Стюарта (8.16), он рассчитал ожидаемую дисперсию для гетерозиготности каждого ферментного локуса, причем параметр θ , или $4N_e\mu$, был оценен по

средней для разных видов гетерозиготности. Ямадзаки вычислил также наблюдаемую дисперсию для каждого фермента у различных видов и сравнил наблюдаемую и ожидаемую вели-

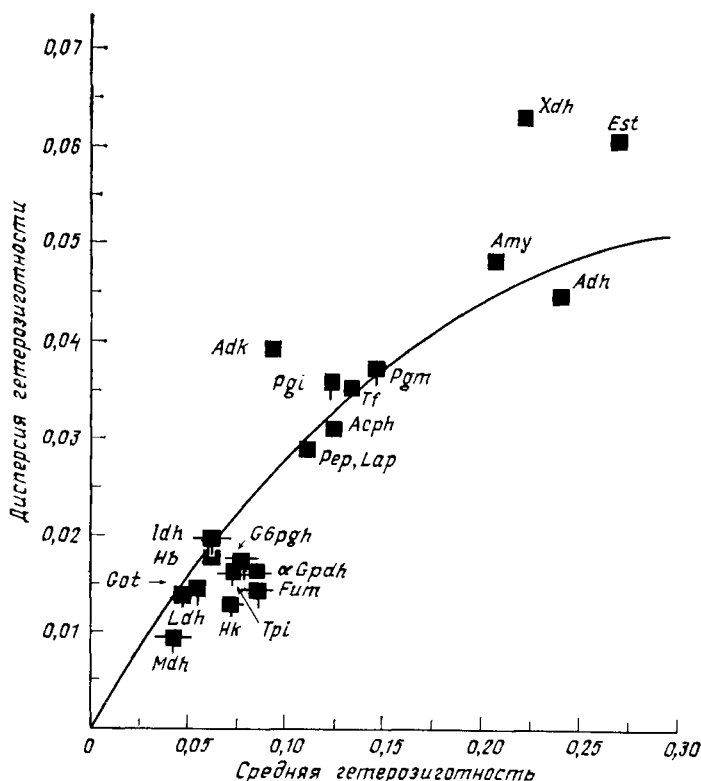


Рис. 9.5. Зависимость между средней гетерозиготностью и дисперсией гетерозиготности для 20 разных белков (в основном ферментов). Наблюдаемые значения средней гетерозиготности представлены на графике квадратами. Кривая отвечает теоретической зависимости, полученной на основе модели бесконечного числа нейтральных аллелей. Черточки, идущие от квадратов, соответствуют различным функциональным ограничениям: черта, направленная вниз, означает, что данный фермент является субстрат-специфичным; черточка, направленная влево, — то, что фермент катализирует единственный метаболический путь; направленная вправо — то, что фермент принимает участие в основных метаболических путях. (По [134], с небольшими изменениями.)

чины по 22 локусам, 13 из которых кодировали субстрат-специфичные, а 9 — субстрат-неспецифичные ферменты. Для 13 субстрат-специфичных ферментов наблюдаемая дисперсия была всегда меньше ожидаемой, причем для 9 ферментов разница между ними была статистически значимой. С другой стороны,

для 9 субстрат-неспецифичных ферментов такой тенденции не отмечалось: наблюдаемая дисперсия была меньше ожидаемой для 5 ферментов и больше для 4, причем различие было статистически значимым лишь в двух случаях.

Аналогичное исследование было выполнено Годжобори [134, 135] с привлечением данных по 20 различным белкам (в основном ферментам) у 14 видов *Drosophila*, 14 видов *Anolis* и 31 вида других организмов. Полученные им результаты приведены на рис. 9.5, где наблюдаемые величины обозначены черными квадратиками. Функциональные ограничения представлены черточками. Черта, направленная вниз, означает, что данный фермент является субстрат-специфичным; черта, направленная влево, — что фермент участвует в одном-единственном метаболическом пути, а черта, направленная вправо, — что фермент участвует в каких-то важных метаболических путях (подробности см. в работе [135]). Кривая на этом рисунке представляет теоретическую зависимость между средней гетерозиготностью и дисперсией гетерозиготности, полученную на основе модели бесконечного числа нейтральных аллелей [см. формулы (8.15) и (8.16) в предыдущей главе]. Из рисунка видно, что для ферментов, имеющих различные функциональные ограничения (они обозначены квадратиками с одной или более черточками), характерен низкий уровень гетерозиготности. Более того, дисперсия гетерозиготности для этих ферментов меньше величины, ожидаемой на основе модели бесконечного числа нейтральных аллелей.

С точки зрения теории нейтральности эти наблюдения лучше всего объясняются тем, что полиморфная изменчивость таких ферментов в основном обусловлена эффективно нейтральными мутациями с очень слабым неблагоприятным действием. Для изучения этого вопроса мной было проведено моделирование методом Монте-Карло популяционных процессов на основе модели эффективно нейтральных мутаций. Как было показано в разд. 8.7, в этой модели предполагается, что селективная невыгодность (s') мутантных аллелей следует гамма-распределению. Был рассмотрен случай, когда $\beta=0,5$ [см. уравнение (8,98)], а популяция состоит из 500 размножающихся особей ($N=N_e=500$). Средний коэффициент отбора ($\bar{s}'$) и скорость мутирования (v) варьировали от эксперимента к эксперименту; величина $\bar{s}'$ изменялась в пределах $0,1 \sim 0,5$, однако в основном я использовал значение $\bar{s}'=0,2$. Предполагаемые скорости мутирования изменялись от 0,00025 до 0,01. Отметим, что средняя дисперсия гетерозиготности определяется не величинами N_e , $\bar{s}'$ и v как таковыми, а их произведениями $N_e\bar{s}'$ и N_ev , поэтому проведенное мной моделирование позволяет судить о поведении больших популяций (какие обычно и встречаются

в природе), даже если при моделировании размер популяции предполагается небольшим ($N_e=500$).

На рис. 9.6 результаты моделирования представлены черными кружками (точками). Например, точка в нижнем левом углу относится к случаю, когда $N_e=500$, $\bar{s}'=0,2$ и $v=0,0004$, откуда $4N_e\bar{s}'=400$ и $4N_ev=0,8$. Отметим, что это эквивалентно ситуации, при которой $N_e=10^6$, $\bar{s}'=10^{-4}$ и $v=2 \cdot 10^{-7}$. Из рисун-

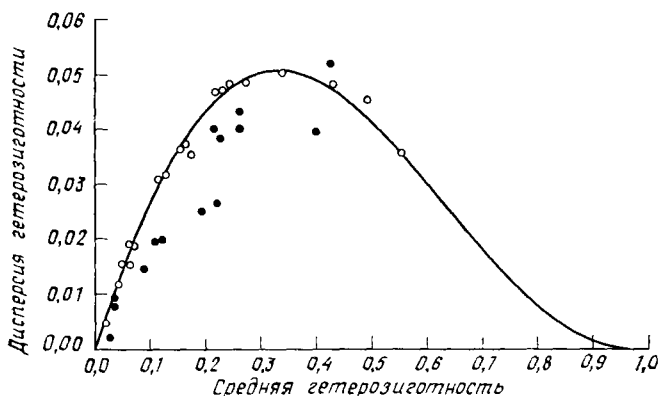


Рис. 9.6. Зависимость между средней гетерозиготностью и дисперсией гетерозиготности, установленная в ходе машинных экспериментов с помощью метода Монте-Карло. Черные кружки — результаты, полученные в рамках модели эффективно нейтральных мутаций при $\beta=0,5$ [см. формулу (8.98)]. Каждая точка отвечает результатам 1000 «наблюдений», сделанных с интервалом в 20 поколений. При моделировании размер популяции предполагается равным 500 ($N_e=N=500$). Средний коэффициент отбора, направленного против мутаций ($\bar{s}'$), считался равным 0,1—0,5, однако в большинстве опытов я принимал его равным 0,2. Скорость мутирования v изменялась в интервале от 0,00025 до 0,01. В каждом эксперименте $\bar{s}'$ и v были постоянными. Светлые кружки отвечают результатам, полученным в предположении полной нейтральности мутаций ($\bar{s}'=0$) и хорошо ложатся на теоретическую кривую, полученную в рамках модели бесконечного числа аллелей. Для случая полной нейтральности мутаций размер популяции при моделировании принимали равным 100 особям, а скорость мутирования варьировали от эксперимента к эксперименту, с тем чтобы получить различные значения средней гетерозиготности.

ка видно, что в том случае, когда полиморфизм обусловлен слабо вредными мутациями, частоты которых следуют гамма-распределению, наблюдаемая дисперсия гетерозиготности должна быть значительно ниже величины, ожидаемой при полной нейтральности (показанной на рисунке кружками и сплошной линией), при условии, что $4N_e\bar{s}'$ достаточно велико (скажем, 400). При увеличении $4N_e\bar{s}'$ смещение наблюдаемых значений вниз становится все более выраженным. Хотя многое

здесь еще не ясно, создается впечатление, что на основе модели эффективно нейтральных мутаций при $N_e\bar{s}=10\sim 100$ можно объяснить наблюдаемую картину гетерозиготности по локусам, кодирующим некоторые ферменты, например гексокиназы и фумаразы, для которых характерны низкий уровень гетерозиготности и жесткие функциональные ограничения.

Как отмечалось несколькими авторами, растворимые ферменты могут подвергаться более существенным мутационным изменениям, чем структурные белки и ферменты, связанные с мембранами. Лэнгридж [293] исследовал влияние замещений аминокислот на ферментативную активность β -галактозидазы *Escherichia coli*. Выделяя мутантов, у которых активность β -галактозидазы не превышала 50% от нормальной, он показал, что только 11 из 733 замещений аминокислот приводили к такому уменьшению активности. Кроме того, в случае β -галактозидазы потеря активности, по-видимому, может частично компенсироваться более интенсивным синтезом этого фермента. Лэнгридж пришел к выводу, что большинство мутаций в генах, кодирующих растворимые белки, в обычных условиях практически нейтральна по отношению к естественному отбору, хотя это и не означает, что какая-то определенная аминокислотная замена будет нейтральной в любых возможных условиях; многие замены проявляются при изменении типа субстрата или его концентрации, при других температуре, pH и т. д.

На структурные белки и белки, связанные с мембранами, обычно налагаются более жесткие функциональные ограничения, чем на растворимые белки. Структурные белки тесно связаны друг с другом, а на мембранные белки налагаются жесткие стерические ограничения, определяющиеся их взаимодействием с другими молекулами, входящими в состав мембран. Поэтому белки этих двух типов в высшей степени инвариантны и характеризуются низким уровнем полиморфизма [195]. Клозе и Феллер [273] исследовали изменчивость локусов, кодирующих мембранные белки и белки цитоплазмы (водорастворимых белков), путем сравнения методом двумерного электрофореза двух инбредных линий мышей. Для белков цитоплазмы была выявлена высокая степень генетической изменчивости (в основном касающейся количества этих белков), однако для белков мембраны генетической изменчивости не наблюдалось.

Ранее Браун и Лэнгли [30] исследовали методом двумерного электрофореза уровень изменчивости белков, присутствующих в больших количествах у взрослых особей *Drosophila melanogaster*, взятых из природной популяции. Средняя гетерозиготность на локус по ~ 54 исследованным локусам составила 4%, что значительно ниже соответствующей величины (14%), полученной для ферментных локусов *D. melanogaster* стандартным методом электрофореза в крахмальном геле. Аналогичное

исследование гетерозиготности таких же белков было проведено в природной популяции домовый мыши. Определенная по ~ 72 структурным локусам средняя гетерозиготность составила 2%, что гораздо ниже соответствующей величины (17,2%), полученной с помощью электрофореза в крахмальном геле.

Более низкий уровень гетерозиготности белков, присутствующих в тканях в больших количествах, по сравнению с растворимыми белками отчасти объясняется тем, что такие белки должны оказывать гораздо более сильное влияние на приспособленность при изменении их структурных свойств в результате мутаций, чем белки, присутствующие в клетке в небольших количествах [см. уравнение (7.11)]. Отметим также, что в больших количествах в клетке присутствуют структурные белки, например актин, тубулин и т. д. [319], для которых, как известно, характерна высокая эволюционная консервативность.

В гл. 7 мы детально рассмотрели, каким образом отрицательный отбор, обусловленный функциональными и структурными ограничениями, налагаемыми на молекулы, сдерживает появление нейтральных мутаций, замедляя таким образом молекулярную эволюцию. Если теория нейтральности верна, те же факторы должны приводить к уменьшению внутривидовой генетической изменчивости. Различные наблюдения, рассмотренные в настоящей главе, полностью согласуются с этим предсказанием. В настоящее время наши представления о природе молекулярных ограничений крайне неполны. Весьма ценным в этом отношении является проведенный в работе [423] детальный физико-химический анализ того, как возникают такие ограничения в результате взаимодействия между боковыми цепями аминокислотных остатков и как они влияют на изменчивость белков, хотя весь этот анализ был проведен для придуманной автором, не существующей в природе молекулы «мембранина». Надеюсь, что скоро такие работы появятся и применительно к реальным белкам, например гемоглобинам, цитохрому *c* и др. (познакомиться с некоторыми новыми результатами в этой области читатели могут в работах [300, 473]).

Из всего сказанного следует, что если теория нейтральности справедлива, то для молекул (или их сегментов), наиболее важных в функциональном отношении и в связи с этим медленно эволюционирующих, будет наблюдаться более низкий уровень гетерозиготности. С другой стороны, в некодирующих областях (особенно в спейсерных участках), а также в интронах (даже у генов, кодирующих высококонсервативные белки) будет присутствовать большое количество «полиморфных» мутаций (не только замен оснований, но также вставок и делеций). Таким образом, фактором, определяющим уровень гетерозиготности, должен быть не балансирующий отбор, а мутационный процесс, размер популяции и отрицательный отбор.

В связи с этим следует упомянуть работу Лэттера [295, 296]. Он проанализировал литературные данные по 25 локусам, кодирующим ферменты, у 9 видов *Drosophila* (5 видов умеренной зоны и 4 тропических вида), причем рассматривал по отдельности белки I и II групп (по классификации [278]). Был рассмотрен вклад аллелей различных частотных классов в гетерозиготность C_i (подробный анализ см. в [296]) в зависимости от частоты аллеля $\bar{q}_i$; эта функциональная зависимость, построенная по результатам наблюдений, сравнивалась с теоретической. Методом машинного моделирования Лэттер показал, что модель, предполагающая возникновение в локусах нейтральных или слабо вредных мутаций, во всех отношениях соответствует имеющимся данным.

Согласно Лэттеру, четкое различие между ферментами групп I и II может объясняться различиями в средней интенсивности отбора. Кроме того, высокосущественное различие в распределении гетерозиготности между тропическими видами и видами умеренной зоны, вероятно, обусловлено меньшим размером популяции последних. Лэттер подчеркнул, что селективная невыгодность мутантов в рамках модели, адекватной экспериментальным данным, была малой: максимальное значение kNs равнялось 3, где kN — суммарный размер размножающейся популяции данного вида (т. е. размер k субпопуляций, каждая из которых имеет эффективный размер N), а s — коэффициент отбора против данной мутации. По мнению Лэттера, с помощью модели, предложенной Кимурой [229] и модифицированной Отой [437, 440, 441] для случая слабонеблагоприятных мутаций, можно объяснить все особенности этих данных. Лэттер считает, что с эволюционной точки зрения значительная часть наблюдаемой в природе электрофоретической изменчивости может с полным основанием считаться селективно нейтральной.

9.6. Распределение частот редких аллелей

Аллели, кодирующие редкие варианты белков, нельзя считать компонентами полиморфных систем, поскольку их частоты слишком низки. Тем не менее изучение таких аллелей может дать ценную информацию относительно механизмов поддержания полиморфизма на молекулярном уровне. Дело в том, что, как было показано в гл. 8 (см., в частности, разд. 8.4), даже явно благоприятная или неблагоприятная в селективном отношении мутация (для которой $|s|$ предполагается малым, однако $4N_e|s|$ может быть большим) ведет себя как селективно нейтральная, если ее частота мала. Отсюда, используя данные по распределению редких аллелей, можно найти скорость мутирования [см. формулу (8.24) и последующее объяснение].

Рассмотрев данные по камбале и японскому макаку, я показал, что в среднем только 1/10 мутаций, связанных с замещением аминокислот, селективно нейтральны, тогда как остальные вредны и не вносят вклада ни в полиморфизм, ни в эволюцию. Здесь я вновь обращаюсь к этой проблеме, используя в основном сведения по популяциям человека, для которых имеются достаточно полные данные.

Харрис и др. [173] оценили частоту встречаемости редких аллелей по 43 локусам, кодирующим ферменты, у европеоидов. Я выбрал из табл. 1 работы [173] 26 локусов, для которых размер выборки превышал 1000. Средний размер выборки для этих локусов $\bar{n}=4058,04$, а среднее число редких аллелей на локус — 49/26, т. е. 1,89. Поскольку по определению Харриса и др. [173] редким считается аллель, частота которого в выборке менее 0,005, в уравнении (8.26), т. е. $M_q = \bar{n}_a(x < q) / \ln(2nq)$, величину q необходимо принять равной 0,005. Тогда, подставляя в это уравнение $\bar{n}_a(x < q) = 1,89$, $\bar{n} = 4058,04$ и $q = 0,005$, мы получим $M_q = 0,509$. Отсюда можно оценить величину $4N_e v_T$, где v_T — полная скорость мутирования на ферментный локус для электрофоретически различимых аллелей. С другой стороны, предположив, что $\bar{H}_e = 0,07$ (приближенное значение средней гетерозиготности на локус для обычных полиморфных аллелей [171]), по формуле (8.20), т. е. $M = \bar{H}_e / (1 - \bar{H}_e)$, получаем $M = 0,0753$. Это позволяет оценить $4N_e v_0$, где v_0 — скорость мутирования для селективно нейтральных аллелей. Таким образом, как следует из (8.27), можно вычислить долю селективно нейтральных мутаций от общего числа мутаций, выявляемых методом электрофореза:

$$P_{\text{нейтр.}} = M/M_q \approx 0,15.$$

Эта величина очень близка к соответствующей оценке, полученной в гл. 8 для камбалы.

Существует и другой метод оценки $4N_e v_T$ (или M_q), который тоже основан на идее, что поведение мутантных аллелей — когда их частота низка — обычно напоминает поведение нейтральных аллелей. Пусть $F_T(x < q)$ — сумма частот (на локус) аллелей, частота каждого из которых меньше q . Тогда в предположении, что размер выборки достаточно велик, т. е. $q \gg 1/(2n)$, мы имеем

$$F_T(x < q) = \int_0^q x \Phi(x) dx \approx 4N_e v_T q = M_q q. \quad (9.3)$$

Отметим, что в отличие от $\Phi(x)$, стремящегося к бесконечности при $x \rightarrow 0$, подынтегральное выражение $x\Phi(x)$ в приведенной выше формуле при $x=0$ принимает конечное значение, поэто-

му мы можем, не допуская большой ошибки, принять в качестве нижнего предела интегрирования 0. Таким образом, $4N_e v_T$ можно определить по формуле

$$M_q' = F_T(x < q)/q, \quad (9.4)$$

где M_q' обозначает определенное этим методом $4N_e v_T$. Согласно [173], средняя гетерозиготность на локус, обусловленная редкими аллелями, равна 1,76 на 1000, так что $F_T(x < q) = 1,76/2000$. Приняв во внимание, что в этом случае $q = 0,005$, получаем $M_q' = 0,176$. Эта величина составляет всего лишь 1/3 оценки $M_q = 0,509$, полученной по формуле (8.26). Поскольку M_q' , определяемое по формуле (9.4), очень сильно зависит от q в знаменателе, вероятность ошибки при оценке M_q' больше, чем при оценке M_q , и поэтому первая из этих величин менее надежна. Таким образом, мы будем считать M_q лучшей оценкой $4N_e v_T$, чем M_q' .

Харрисом и др. [173] сделано важное замечание относительно того, что средняя гетерозиготность по редким аллелям «полиморфных» и «мономорфных» локусов не различается, если исключить из рассмотрения щелочную фосфатазу плаценты, для которой характерна необычайно высокая изменчивость. С точки зрения теории нейтральности это легко понять: истинная скорость мутирования (v_T) у этих двух типов локусов практически одинакова, поэтому, как следует из уравнения (9.3), значения $F_T(x < q)$ для них должны быть приблизительно одинаковыми.

Интересные данные получены Нилом и его сотрудниками в результате обширных исследований редких вариантов белков в популяциях американских индейцев. Я проанализировал данные, представленные в табл. 1 работы Нила [398], где приведена частота редких вариантов по 28 локусам в 21 трибе американских индейцев. В соответствии с тем как Нил определяет редкие аллели, $q = 0,01$ по нашей терминологии, и из его таблицы мы получаем $n_a(x < q) = 1,29$, $n = 6442,07$, откуда $M_q = 0,266$. Отметим одну интересную особенность данных Нила: для некоторых вариантов наблюдается так называемый «родственный полиморфизм» (по терминологии Нила): они концентрируются в одном или нескольких родственных племенах, где их частоты значительно превышают минимальную для данного полиморфизма частоту. Например, аллель YAN-2 локуса, кодирующего альбумин, встречается в племени Яномама с частотой, превышающей 6%, и полностью отсутствует в других племенах. Величину M можно вычислить, используя среднюю гетерозиготность $\bar{H}_e = 0,054$ по 23 локусам из 12 племен, данные по которым приведены в табл. 5 работы [398]. Отсюда получаем $M = 0,0571$. Таким образом, доля нейтральных мутаций $P_{\text{нейтр}}$,

определенная как M/M_q , равна 0,21, что не слишком сильно отличается от соответствующей величины, полученной для популяций европеоидов $P_{\text{нейтр}}=0,15$).

Кстати, эти значения $P_{\text{нейтр}}$ согласуются с результатами, полученными при изучении скорости эволюции псевдогенов, которые мы обсуждали в разд. 7.4. Если использовать оценки, приведенные в табл. 3 работы Ли и др. [312], то средняя скорость замещений нуклеотидов, приводящих к замене аминокислот в нормальных глобиновых генах, составит около 0,14 от соответствующей величины для глобиновых псевдогенов. С точки зрения теории нейтральности последняя оценка (т. е. 0,14) должна представлять собой долю нейтральных мутаций от общего числа мутаций в генах глобинов. Учитывая тот факт, что скорости эволюции глобинов близки к медиане скорости эволюции белков (гл. 4), это соответствие можно рассматривать как еще одно свидетельство в пользу теории нейтральности.

Аналогичным образом можно проанализировать данные по пяти видам дрозофилы группы *willistoni*, исследованным Айалой и др. [13], хотя здесь размер выборки на локус на вид меньше, чем для данных по популяциям человека. Из всех исследованных 5 видов только у *D. willistoni* средний размер выборки на локус (выраженный через число генов, т. е. $2n$) превышал 500, поэтому именно на нем я и остановлюсь. Из табл. 1 работы [13], где приведены аллельные частоты для 13 локусов, я выбрал те аллели, частота которых составляла менее 1% ($q=0,01$). Таких аллелей было 85, т. е. $\bar{n}_a(x < q) = 85/31 \approx 2,74$. Средний размер выборки на локус равен $2\bar{n} = 568,06$. Отсюда получаем $M_q = 1,60$. Значение средней гетерозиготности на локус, приведенное в табл. 6 работы [13], равно 0,177, откуда $M = 0,215$. Таким образом, наша оценка доли нейтральных мутаций среди всех электрофоретически различных мутаций в момент возникновения $P_{\text{нейтр}} = M/M_q = 0,13$. Интересно отметить, что эта оценка не слишком сильно отличается от соответствующих оценок, полученных для популяций человека.

Вероятно, Ота [438] была первой, кто отметил, что, хотя наблюдаемые и теоретические распределения частот полиморфных аллелей в рамках теории нейтральности довольно хорошо соответствуют друг другу, в наблюдаемом распределении имеется заметный избыток редких аллелей. Проведенный нами анализ подтверждает ее результаты. Далее Ота показала [439, 440], что для субстрат-специфичных ферментов этот избыток больше, чем для субстрат-неспецифичных. Действительно, если применить рассмотренный выше метод к каждой из этих групп ферментов по отдельности, мы получим $P_{\text{нейтр}} = 0,070$ для субстрат-специфичной группы и 0,204 — для субстрат-неспецифичной, что согласуется с данными Оты. Полученные значения

$P_{\text{нейтр}}$ свидетельствуют о том, что вероятность нейтральности мутационной замены гораздо меньше для субстрат-специфичных ферментов по сравнению с субстрат-неспецифичными. Это свидетельствует о правильности объяснения с позиции теории нейтральности того факта, что гетерозиготность на локус для субстрат-специфичных ферментов гораздо ниже, чем для субстрат-неспецифичных (см. предыдущую главу).

Ясно, что подробное исследование аллелей, кодирующих редкие варианты, так же важно для понимания механизмов поддержания генетической изменчивости, как и изучение полиморфных аллелей. Надеюсь, что в отношении редких аллелей тех природных видов, экология которых (в особенности размер популяций) хорошо известна, будет получено больше данных.

9.7. Нейтральные аллели в географически подразделенной популяции

Очень часто заключение о типе отбора делается на основе наблюдаемой картины полиморфизма в природной популяции. Так, если обнаруживается, что два аллеля имеют одинаковые частоты по всему ареалу данного вида, делается вывод, что эти частоты активно поддерживаются естественным отбором. С другой стороны, если в различных локальных популяциях закреплены разные аллели или если обнаружена клина частот, это нередко считается результатом локальной адаптации данных аллелей. Более того, если частоты аллелей одинаковы в пределах локальности, но различаются в разных локальностях, предполагается, что это указывает на существование какого-либо вида «балансирующего отбора».

Вообще говоря, любую картину полиморфизма в природных популяциях можно объяснить действием естественного отбора. Часто наличие такого предполагаемого отбора используется как свидетельство против теории нейтральности. Однако, как указывали Кимура и Маруяма [251], те же самые распределения частот могут быть объяснены исходя из миграции и случайного дрейфа селективно нейтральных мутаций.

В предыдущей главе было показано (см. разд. 8.3), что если в двумерной модели «степпинг-стоун» предполагается миграция по крайней мере нескольких особей между соседними колониями (демами) в каждом поколении, то частоты полиморфных аллелей в разных колониях будут практически одинаковыми. Точнее, как показано Маруямой [332] на основе математического анализа двумерной модели «степпинг-стоун», если N — эффективный размер каждой колонии, а m — скорость, с которой данная колония обменивается особями с четырьмя

соседними колониями, то при

$$mN \geq 4 \quad (9.5)$$

популяция является практически панмиктической.

Вообще говоря, для того чтобы распределение аллельных частот в колониях было равномерным, необходимо выполнение дополнительного условия

$$N_T v \ll 1, \quad (9.6)$$

где N_T — эффективный размер всей популяции, а v — скорость мутирования (см. [251]). Для большинства организмов это условие в отношении полиморфизма белков, очевидно, выполняется, поскольку значение средней гетерозиготности редко превышает 30%. С другой стороны, если

$$mN < 1, \quad (9.7)$$

может возникнуть заметная локальная дифференциация генных частот.

Хорошо известно, что у *Drosophila* частоты полиморфных аллелей в локальных популяциях одного вида обычно сходны (см., например, [13]). Кимура и Ота [256] объяснили этот факт с точки зрения теории нейтральности, отметив, что для видов *Drosophila* условие (9.5) легко выполняется из-за их высокой миграционной способности. Айала и др. [13], возражая против этого объяснения, отмечают, что если весь вид практически панмиктичен, то размер популяции некоторых неотропических видов *Drosophila* (например, *D. willistoni*) должен быть крайне большим — по крайней мере 10^9 , а возможно, даже 10^{10} или более. Поэтому Айала и др. утверждают, что даже при такой низкой скорости мутирования, как $v = 10^{-7}$, величина $4N_T v + 1$ становится равной 401, и, следовательно, ожидаемая гетерозиготность составляет почти 100%, что противоречит наблюдаемым фактам. Этот вопрос я уже рассматривал в разд. 9.1 и не буду обсуждать его здесь еще раз.

С тех же позиций, но с привлечением более детального математического анализа подвергает критике теорию нейтральности Балмер [35], указывая, что гипотеза о бесконечном числе нейтральных аллелей [246] совместно с моделью двумерной миграции [326] ведет к результатам, не соответствующим реальным данным. Балмер отмечает следующие два противоречия. Во-первых, из формулы Малекó, использующейся в модели непрерывной двумерной миграции, следует довольно четко выраженная географическая локализация аллельных частот, тогда как реальные данные свидетельствуют об их равномерном распределении. Балмер применил два параметра из модели Малекó: коэффициент инбридинга f_0 , т. е. вероятность того, что два гомологичных локуса данной особи идентичны по про-

исхождению, и коэффициент родства f_x , определенный аналогичным образом для двух особей, разделенных расстоянием x . Используя формулы Малекó для f_0 и f_x , Балмер вычислил отношение $\rho_x = f_x/f_0$. Приведенные им численные значения ρ_x очень быстро уменьшаются с увеличением x . С другой стороны, оценки этого отношения по данным для *Drosophila pseudoobscura* [473] очень близки к единице даже для весьма отдаленных друг от друга местностей (например, для Техаса, Колорадо и Калифорнии). Балмер считает, что такое несоответствие нельзя объяснить, если не принять абсурдного предположения, что корень из среднего квадрата миграционных расстояний (σ) имеет величину порядка 2000 км.

Во-вторых, Балмер отмечает, что ожидаемый уровень гомозиготности f_0 намного ниже соответствующей экспериментальной величины, за исключением тех случаев, когда миграция резко ограничена. Значение средней гомозиготности, полученное Пракашем и др. [473], составляет 0,877. Используя формулу Малекó для определения f_0 , Балмер утверждает, что для того, чтобы f_0 была такого порядка величины, средний квадрат миграционных расстояний между местами размножения родителей и потомков (σ^2) должен быть неправдоподобно малым. Например, если скорость мутирования $v = 10^{-5}$, то число особей в круге радиуса σ , т. е. $\pi\sigma^2\delta$ (где δ — плотность размножающейся популяции), будет равно примерно 0,38.

Выступая в защиту теории нейтральности, Маруяма и Кимура [339] подчеркнули следующий важный факт: формулы Малекó, использованные Балмером, справедливы только для случая, когда общая численность популяции N_T гораздо больше величины, обратной скорости мутирования (v), а ареал достаточно велик. Однако размер ареала сам по себе не является столь существенным, если популяция имеет конечную численность. Безусловно, для популяции бесконечного размера формулы Малекó справедливы. С другой стороны, если N_T и $1/v$ — величины одного порядка, необходимо пользоваться формулами, справедливыми для конечной популяции. В этом случае формулы Малекó непригодны даже как приближенные, поскольку они дают совершенно другие значения. Точные решения для конечного тороидального пространства размером $L_1 \times L_2$ были получены Маруямой [332]. Хотя для более реалистичного прямоугольного ареала размером $L_1 \times L_2$ точных решений не найдено, приближенные решения для этого случая можно получить, используя тороидальное пространство размером $2L_1 \times 2L_2$. Приняв $L_1 = L_2$, Маруяма и Кимура [339] выполнили на ЭВМ численные расчеты большого объема.

В отличие от Балмера [35], Маруяма и Кимура пришли к выводу, что теоретические и экспериментальные данные соответствуют друг другу. Если $N_T v \leq 1$, а $\sigma^2\delta$ имеет порядок 100,

10 или даже менее, отношение $f(x, y)/f_0$ остается очень близким к 1 и почти не зависит от расстояния и размера ареала $[f(x, y)$ — вероятность того, что два гомологичных гена, разделенных расстояниями x и y по первой и второй осям декартовой системы координат, идентичны по происхождению]. Далее, если $4N\tau v = 0,1 \sim 0,2$, то f_0 по порядку величины приблизительно соответствует предположениям теории нейтральности. Как считают Маруяма и Кимура [339], вследствие ограниченности численности популяции ситуация становится совершенно иной, чем в случае, когда размер популяции бесконечен, и, таким образом, Балмер необоснованно использует формулы Малекб, выведенные для популяций бесконечного размера. Это и приводит его к ошибочным заключениям, казалось бы не согласующимся с наблюдаемыми данными. На самом же деле дилемма Балмера разрешается просто путем использования более подходящих формул.

Бейкер [16] исследовал популяцию белоголовой воробьиной овсянки (*Zonotrichia leucophrys nuttalli*) с определенным «певческим диалектом», обитающую вдоль центрального побережья Калифорнии. Эта популяция состоит примерно из 2000 птиц; «величина соседства» (neighborhood size) по терминологии Райта [596], определенная Бейкером, составляет $N=36$, т. е. около 100 особей. По наблюдениям автора [16], каждая группа особей этого вида с определенным «диалектом» в значительной мере изолирована от других групп; оценка F -статистики Райта F_{ST} из данных по наблюдаемым внутри группы частотам изотипов дает среднее значение $F_{ST}=0,0046$. Бейкер вычислил также теоретическое значение F_{ST} , используя модель изоляции расстоянием Райта [599], и получил величину на порядок выше наблюдаемой (например, для 20 «соседств», каждое из которых имеет размер 36, $F_{ST}=0,0475$). По мнению Бейкера, в пределах данной диалектной группы отбор, по-видимому, поддерживает большую гомогенность, чем ожидается исходя из предположения о нейтральности аллелей. Однако он также отметил, что в рамках тороидальной модели Маруямы [335] при скорости мутирования $v=2 \cdot 10^{-6}$ ожидаемые значения F_{ST} близки к наблюдаемым; это предполагает отсутствие значительных отклонений от панмиксии внутри диалектных популяций.

Мне хотелось бы закончить данный раздел, приведя еще несколько примеров, свидетельствующих о том, что изоляция и случайный дрейф являются основными факторами в определении географического распределения частот аллелей. Сакаидзуми и др. [499] исследовали географическое распределение аллелей 25 ферментных локусов у *Oryzias latipes* — пресноводной рыбки, обитающей в Японии в ручьях и на рисовых полях. Для пяти из 25 исследованных локусов обнаружена четкая дифференциация аллельного состава между популяциями вдоль по-

бережья Японского моря, с одной стороны, и популяциями вдоль тихоокеанского побережья — с другой. На рис. 9.7 показано распределение частот аллелей, кодирующих алкогольдегидрогеназу (*Adh*) и эстеразу-1 (*Est-1*). Характер этого распределения можно объяснить, заметив, что две рассматриваемые группы популяций изолированы друг от друга горным хребтом, разделяющим территорию Японии на две части. Ясно, что горный хребет с давних пор является для этой рыбки не-

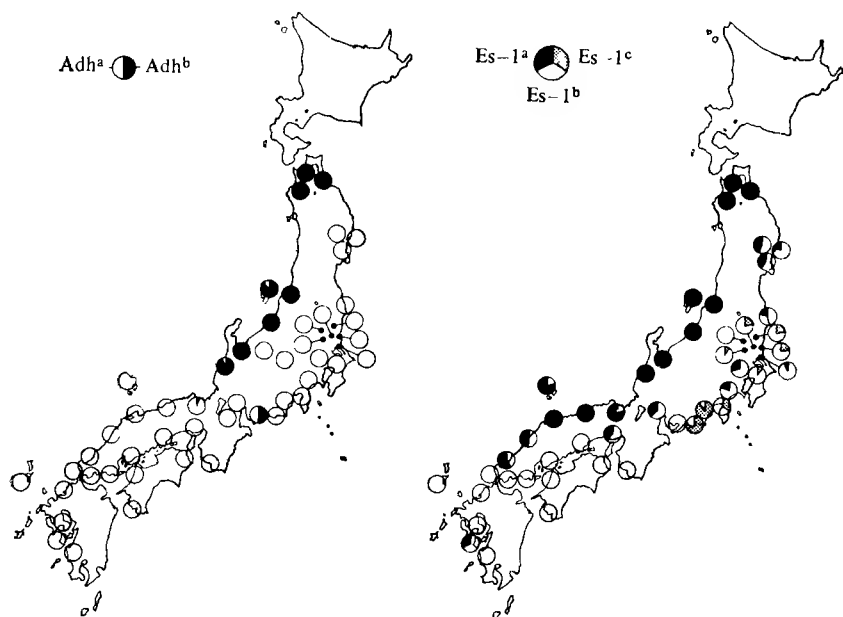


Рис. 9.7. Географическое распределение аллелей локуса, кодирующего алкогольдегидрогеназу (слева) и эстеразу-1 (справа), у *Oryzias latipes*. (По [498], с небольшими изменениями.)

преодолимым препятствием, в связи с чем по данным пяти локусов произошло (или почти произошло) случайное закрепление нейтральных аллелей.

Совершенно другая картина — отсутствие дифференциации аллельного состава между популяциями двух прибрежных районов — наблюдается у подовой мушки *Drosophila virilis*, изученной Обой [428]. Оба определил в больших выборках из этих популяций аллельные частоты 7 ферментных локусов, включая *Est-α*, *Est-β*, *Adh* и т. д. Полученные результаты легко объяснить, вспомнив, что дрозофила способна активно мигрировать, поэтому ни одна ее локальная популяция не изолирована от других в достаточной степени.

В заключение я хотел бы рассмотреть случай, когда условие (9.6) не выполняется. Несколько лет назад Робертсон [489] отметил, что в случае нейтральности значительной доли мутаций по данному локусу могут иметь место две ситуации: либо расщепление очень небольшого числа аллелей в больших популяциях, либо наличие небольшого числа различных наборов аллелей в изолированных популяциях малого размера. По мнению Робертсона, поскольку ни одна из этих альтернативных ситуаций не наблюдалась, большая часть полиморфизма активно поддерживается или поддерживалась естественным отбором. Кимура и Ота [256] указали, что обе отмеченные Робертсоном ситуации характерны для полиморфизма распределения гетерохроматина в хромосомах (выявляемого обработкой холодом) в природных популяциях двулетнего растения *Trillium kamtschaticum*. Обширные цитологические исследования этого растения [151, 152] показали, что в большой популяции наблюдается расщепление нескольких типов хромосом, тогда как в малых изолированных популяциях закреплены различные хромосомные типы. Действительно, предположение Робертсона справедливо, если субпопуляции почти полностью изолированы и если скорость мутирования в случае нейтральной изменчивости достаточно высока для появления в популяции более чем одного мутанта в каждом поколении. Полиморфизм хромосом *Trillium* можно объяснить, допустив сравнительно высокую скорость мутирования на хромосому, поддерживающую гетерохроматиновую изменчивость, и предположив, что скорость миграции на поколение очень низка.

Заключение

Традиционные исследования эволюции ограничивались изучением только внешних признаков (например, крыльев птиц, шеи жирафы, зубов лошади и т. д.) и их функций. В основе любых объяснений эволюционных изменений фенотипов лежала дарвиновская теория естественного отбора, или принцип выживания наиболее приспособленного. Во времена Дарвина механизм наследования и природа наследственной изменчивости были неизвестны, и в связи с этим Дарвин испытывал определенные трудности, однако благодаря своей глубокой интуиции он сумел осознать действительное значение естественного отбора в эволюции.

С возникновением менделевской генетики трудности, с которыми пришлось столкнуться Дарвину, были постепенно разрешены, а развитие популяционной генетики привело к объединению (синтезу) дарвинизма и менделизма. Так появилась «синтетическая теория эволюции», которая, бурно развиваясь, к концу 50-х годов превратилась во впечатляющую систему взглядов. Хотя эта теория рассматривает различные факторы эволюции — например, мутационный процесс, рекомбинацию, миграцию, — ее наиболее характерной особенностью является то, что главная роль в ней отводилась и отводится естественному отбору. Синтетическая теория утверждает, что скорость и направление эволюции в основном определяются действием направленного отбора (positive selection), а мутационный процесс носит второстепенный характер. Именно поэтому данная теория получила также название теории неodarвинизма.

К концу 60-х годов среди биологов утвердилось мнение, согласно которому судьбу любого биологического признака можно объяснить в свете адаптивной эволюции как результат действия естественного отбора. Кроме того, считалось, что мутации почти никогда не бывают селективно нейтральными, а случайный дрейф генов не имеет для эволюции существенного значения (исключение составляет теория смещающегося равновесия Райта). Исходя из этих теоретических положений и основываясь на разнообразных исследованиях — генетическом ана-

лизе природных популяций, палеонтологических исследованиях, экологической генетике и изучении видообразования, — было написано множество работ, рассматривающих формирование генофондов видов и их изменение в ходе эволюции. Казалось, что механизм эволюции можно объяснить только с позиций неodarвинизма.

На самом деле все эти заключения основывались на фенотипических наблюдениях, и узнать, что же происходит при эволюции на молекулярном уровне, т. е. на уровне внутренней структуры гена, не представлялось возможным. В то время при объяснении генетической структуры природных популяций важная роль приписывалась эпистазу и сверхдоминированию, о степени же внутривидовой молекулярной изменчивости было известно очень мало.

С возникновением молекулярной генетики и внедрением ее методов и понятий в изучение эволюции ситуация изменилась. Путем сравнительного анализа аминокислотных последовательностей гомологичных белков у родственных организмов (например, сравнения гемоглобинов позвоночных) и с привлечением палеонтологических данных появилась возможность определить скорость и характер замещений аминокислот в эволюции. Это позволило оценить темпы нуклеотидных замен в генах и экстраполировать их на все количество ДНК, присутствующей в рассматриваемом организме. Полученные таким образом факты показали, что скорость накопления мутаций гораздо выше, чем считалось раньше. С разработкой новых методов электрофореза стало возможным быстрое определение изменчивости на уровне отдельных особей. Эти исследования выявили у многих организмов гораздо более высокий уровень генетического полиморфизма, чем это предполагалось раньше.

Благодаря всем этим исследованиям появилось много интересных данных и в то же время возник ряд трудных вопросов. Я попытался ответить на них на основе стохастической теории популяционной генетики, выдвинув свою теорию нейтральной эволюции (или, точнее, гипотезу взаимодействия нейтральных мутаций и генетического дрейфа). В отличие от традиционной синтетической теории эволюции (неodarвинизма) теория нейтральности утверждает, что большая часть мутационных замещений в ходе эволюции обусловлена не положительным дарвиновским отбором, а случайным закреплением нейтральных или почти нейтральных мутаций. Эта теория также предполагает, что в основном внутривидовая молекулярная генетическая изменчивость, проявляющаяся в виде полиморфизма белков, селективно нейтральна или почти нейтральна и поддерживается в популяциях любого вида благодаря равновесию между мутационным процессом и случайной элиминацией или фиксацией аллелей.

С точки зрения теории нейтральности мутационные замещения, осуществляющиеся в ходе эволюции, и молекулярный полиморфизм — это не независимые друг от друга явления, а две стороны одного и того же явления. Другими словами, полиморфизм белков представляет собой просто переходную фазу молекулярной эволюции. Теория нейтральности хорошо обоснована математически: в ее рамках предпринимаются попытки количественно рассмотреть различные вопросы молекулярной эволюции и полиморфизма с точки зрения генетики популяций.

Теория нейтральной эволюции вызвала резкие возражения и была подвергнута критике со стороны «селекционистов», придерживающихся традиционной синтетической теории. По их мнению, мутационные замещения в эволюции адаптивны и обусловлены положительным дарвиновским отбором. Сторонники селекционистской теории считают адаптивным также и белковый полиморфизм и утверждают, что он поддерживается на уровне вида благодаря действию какой-либо формы балансирующего отбора. Другими словами, для объяснения этих двух явлений они привлекают два различных вида естественного отбора. При этом многие биологи сомневаются в состоятельности теории нейтральности, а огромное большинство пока не убеждено в ее корректности.

Некоторые возражения против теории нейтральности связаны с непониманием ее сути. Например, эта теория вовсе не утверждает, что абсолютно все мутации в момент их появления селективно нейтральны (эквивалентны). Напротив, она предполагает, что определенная их часть обладает неблагоприятным эффектом, тогда как остальные селективно нейтральны, причем доля вредных мутаций зависит от функциональных ограничений, налагаемых на молекулу: чем они жестче, тем эта доля больше, а вероятность того, что мутация окажется селективно нейтральной, — ниже. Несомненно, в ходе эволюции возникают и благоприятные мутации, однако в теории нейтральности предполагается, что они слишком редки для того, чтобы иметь значение для объяснения основных особенностей молекулярной эволюции, обнаруженных путем сравнительного изучения нуклеотидных и аминокислотных последовательностей ДНК и белков, а также внутривидовой генетической изменчивости, выявляемой методом электрофореза или другими тонкими методами.

В этой книге я стремился показать, что теория нейтральной эволюции не только успешно выдержала различного рода проверки, направленные на ее опровержение, но и получила неожиданное подтверждение благодаря появлению чрезвычайно перспективной методики клонирования фрагментов ДНК и быстрого определения их нуклеотидных последовательностей.

Мне хотелось бы привести некоторые из этих данных. Было с очевидностью показано, что синонимичные замены (т. е. замены, не приводящие к замещению аминокислот) в тех участках генома, которые кодируют белки, возникают в эволюции гораздо чаще, чем несинонимичные; иногда этот эффект выражен необыкновенно сильно, как в случае гистона H4 и тубулина. Поскольку естественный отбор действует на фенотипы, а для них структура и функция белков играют решающую роль, ожидается, что при прочих равных условиях мутации, не приводящие к замещению аминокислот, будут подвергаться меньшему давлению отбора по сравнению с мутациями, сопровождающимися такими замещениями. Действительно, за последние несколько лет было неопровержимо доказано, что синонимичные замены, а также «молчащие» замены в некодирующих участках ДНК являются преобладающими эволюционными изменениями на молекулярном уровне. Эти данные весьма сходны с соответствующими данными по заменам аминокислот в белках. Например, при образовании из проинсулина активного инсулина средняя часть молекулы проинсулина (пептид С) выщепляется и в дальнейшем не используется. Известно, что скорость эволюции пептида С, выраженная через число аминокислотных замен, в несколько раз выше, чем у инсулина.

Не так давно было показано, что большинство генов эукариот содержат «интроны», которые выщепляются при образовании зрелой РНК и, следовательно, не участвуют в синтезе белка. Как мы видели, эволюционные замещения нуклеотидов происходят в интронах с очень высокой частотой. Еще интереснее тот факт, что и у «псевдогенов», т. е. генов, гомологичных соответствующим нормальным генам, но утратившим свою функцию, нуклеотидные замещения в ходе эволюции тоже происходят с очень высокой скоростью — примерно равной или даже немного большей, чем скорость синонимичных замен. Эти факты указывают на следующую общую закономерность: *молекулярные изменения, которые с меньшей вероятностью подвергаются давлению естественного отбора, возникают в эволюции с большей скоростью.* Эта эмпирическая закономерность легко объяснима с точки зрения теории нейтральности, поскольку у таких изменений больше шансов быть селективно нейтральными (т. е. селективно эквивалентными), и, следовательно, в данном случае нейтральная эволюция в результате генетического дрейфа идет более быстрыми темпами. Как уже упоминалось, теория нейтральной эволюции предполагает, что некоторая часть мутаций — в зависимости от функциональных ограничений — обладает неблагоприятным эффектом. В случае синонимичных замен или нуклеотидных замещений в псевдогенах такие ограничения должны отсутствовать или быть очень слабыми, поэтому здесь большинство (практически все) мута-

ций являются нейтральными. Согласно теории нейтральности, чем слабее функциональные ограничения, тем больше доля селективно нейтральных мутаций и, следовательно, тем выше скорость эволюции. С другой стороны, эта эмпирическая закономерность с очевидностью противоречит интерпретации с позиции неodarвинизма, поскольку для быстрого накопления мутаций требуется сильное давление отбора, благоприятствующего этим мутациям. Единственное объяснение, которое может быть дано с позиции неodarвинизма, состоит в том, что эти быстро эволюционирующие участки выполняют какие-то неизвестные функции и что они подвержены быстрой адаптивной эволюции, происходящей путем накопления благоприятных мутаций.

Эволюционные данные со всей полнотой показали, что для изменений на фенотипическом уровне характерны адаптация и широкое разнообразие путей приспособления, неравномерность скорости в пределах филума и разная скорость в разных филумах. Напротив, при исследовании молекулярной эволюции обнаружили совершенно другие особенности эволюционного процесса, характеризующиеся консервативными по природе и случайными по характеру измерениями, происходящими равномерно, причем скорость эволюции данного белка одинакова в разных филумах.

Последняя особенность, а именно примерное постоянство скорости эволюции, выраженной через число аминокислотных замен, стала одним из наиболее противоречивых вопросов молекулярной эволюции. Особенно интересна эта особенность в связи с тем, что постоянство скорости эволюции проявляется при ее измерении в расчете на год, а не на поколение. Реакция на это открытие у ученых-эволюционистов была разной. Некоторые из них упорно отрицают сам факт постоянства скорости замещения аминокислот. Другие исследователи, напротив, твердо верят в то, что скорость замещения аминокислот строго постоянна, и считают это аксиомой, однако уделяют мало внимания рассмотрению механизма данного явления.

В этой книге я попытался показать, что постоянство скорости эволюции каждого данного белка действительно имеет место, хотя и является приблизительным. Для многих белков, в частности для глобинов, наблюдаемая дисперсия скорости эволюции между филумами (в основном среди млекопитающих) для данного белка намного выше, чем это ожидается исходя из чисто вероятностных соображений. Например, для миоглобина млекопитающих (включая человека, тюленя, барсука, лошадь, крупный рогатый скот и кашалота) отношение наблюдаемой дисперсии к ожидаемой (R) составляет 1,7. В этом случае отклонение от предполагаемого постоянства скорости статистически недостоверно. В других случаях, как мы

видели, значения R варьируют от 1,3 до 3,3. Интересно, что аминокислотные замены происходят в эволюции с одинаковой скоростью в таких филумах млекопитающих, условия и образ жизни которых совершенно различны. Такое постоянство с точки зрения теории нейтральности можно объяснить тем, что для данного белка скорость возникновения нейтральных мутаций в год одинакова в разных филумах.

Это объяснение основано на следующей теореме, применимой к нейтральным мутациям: скорость k , с которой мутантные аллели в данном локусе (или сайте) замещаются в процессе эволюции вида, равна скорости мутационного процесса v . Другими словами, если мы ограничимся рассмотрением нейтральных мутаций, то $k=v$ (при условии, что обе эти величины измеряются в расчете на одну единицу времени). С другой стороны, если мутационные замещения испытывают давление естественного отбора, поддерживающего явно благоприятные мутации, то можно показать, что $k=4N_e\bar{s}_A v_A$, где N_e — эффективный размер популяции вида, v_A — скорость мутирования для благоприятных аллелей, а $\bar{s}_A$ — средний коэффициент отбора для этих аллелей. В данном случае для того, чтобы величина k оставалась постоянной, произведение $N_e\bar{s}_A v_A$ должно быть постоянным в ходе эволюции разных филумов. Поскольку каждая из величин — N_e , $\bar{s}_A$ и v_A — может быть совершенно разной для таких организмов, как человек, тюлень, лошадь, крупный рогатый скот и кашалот, вряд ли их произведение будет одинаковым для всех этих организмов. Таким образом, явление постоянства скорости эволюции трудно объяснить с позиций селекционизма. Иногда утверждается, что его можно объяснить, предполагая одинаковое «давление отбора» в различных филумах, однако это объяснение совершенно неверно — постоянство $\bar{s}_A$ не обеспечивает постоянства k . Более того, в каждом поколении основной формой естественного отбора является стабилизирующий отбор, который не имеет никакого отношения к скорости адаптивных мутационных замен.

Проблема постоянства скорости молекулярной эволюции усложняется, если включить в рассмотрение организмы с совершенно разной продолжительностью поколения, например дрозофилу, мышь и человека. Соблюдается ли правило постоянства скорости эволюции в этом случае — в точности не известно, однако не исключено, что здесь оно носит приближенный характер. Это дало повод некоторым ученым подвергнуть критике теорию нейтральности: исследования по мутагенезу свидетельствуют о том, что скорость возникновения спонтанных мутаций в расчете на поколение (но не на год), вероятно, примерно одинакова у животных с совершенно разной продолжительностью поколения; если теория нейтральности справед-

лива, то из теоремы для нейтральных мутаций следует, что скорость эволюции должна быть постоянна в расчете не на год, а на поколение, а это противоречит действительности.

Я должен признать, что с помощью теории нейтральности решить эту проблему трудно, однако решить ее с позиций положительного отбора еще труднее. В рамках теории нейтральности существуют по крайней мере два подхода к этой проблеме. Можно предположить, что скорость возникновения нейтральных мутаций пропорциональна времени в годах, даже если скорость возникновения видимых и летальных мутаций в расчете на поколение сравнима у организмов с разной продолжительностью поколения. В свете результатов исследования псевдогенов это предположение кажется мне неудовлетворительным, поскольку наблюдаемые у них необыкновенно высокие скорости эволюции можно объяснить исходя из предположения, что мутации, бывшие явно вредными, превратились в нейтральные после того, как данные гены утратили свою функцию и стали псевдогенами. Однако скорость эволюции псевдогенов, вероятно, не пропорциональна продолжительности поколения, хотя для окончательного выяснения этого вопроса нужны дополнительные данные.

Другое объяснение постоянства скорости эволюции в расчете на год может быть основано на гипотезе Оты о мутациях с очень слабым неблагоприятным эффектом. Согласно этой гипотезе мутации, которые мы называем нейтральными, на самом деле имеют очень слабый неблагоприятный эффект, причем величина коэффициента отбора s' сравнима со скоростью мутирования или немного выше. Такие очень слабо вредные мутации ведут себя как селективно нейтральные и в сравнительно небольших популяциях сильно подвержены влиянию генетического дрейфа, тогда как в очень больших популяциях против них действует эффективный отбор. С такой точки зрения истинно нейтральные мутации представляют собой крайний случай селективно неблагоприятных мутаций, когда их вредный эффект становится пренебрежимо малым, а не крайний случай благоприятных мутаций. Предположим, что частотное распределение коэффициентов отбора, направленного против вновь возникающих мутаций, очень асимметрично и сдвинуто в сторону малых значений s' . Тогда у вида с небольшим размером популяции значительная часть мутаций является эффективно нейтральной. С увеличением размера популяции все большее число мутаций начинает испытывать направленное против них давление отбора, а доля нейтральных мутаций уменьшается. Существует следующая общая закономерность: для крупных животных характерна большая продолжительность поколения и в то же время малая эффективная численность популяции. Это значит, что у животных с большей продолжительностью поко-

ления размер популяции обычно меньше, поэтому скорость возникновения мутаций, ведущих себя как селективно нейтральные, у них выше. Следовательно, скорость мутирования для эффективно нейтральных аллелей в расчете на год оказывается примерно одинаковой у разных видов животных. Моя модель «эффективно нейтральных мутаций», в которой предполагается, что величины s' следуют гамма-распределению, является развитием и количественным выражением идеи Оты. В такой новой формулировке скорость эволюции в расчете на поколение равна скорости возникновения «эффективно нейтральных мутаций». В частности, мной рассмотрена ситуация, когда параметр $\beta=0,5$. В этом случае скорость эволюции в расчете на поколение k_g обратно пропорциональна квадратному корню из эффективной численности популяции, т. е. $k_g \propto 1/\sqrt{N_e}$. Если продолжительность поколения (g) обратно пропорциональна корню квадратному из эффективной численности популяции, то скорость эволюции в расчете на год $k_1 = k_g/g \propto (1/\sqrt{N_e})/(1/\sqrt{N_e}) = 1$, т. е. постоянна, если постоянна скорость мутирования на сайт на поколение. Справедливо это предположение или нет, покажет будущее. Отметим, что N_e здесь обозначает эффективную численность популяции за все то длительное время, которое осуществляется эволюция, поскольку скорость мутационных замещений измеряется в расчете на десятки миллионов лет. Следовательно, проверка этого предположения, основанная на данных по дрозофиле с использованием наблюдаемых в настоящее время эффективных размеров популяций различных видов этого организма, не может быть признана адекватной. Дело здесь не только в том, что для дрозофилы нет надежных оценок скоростей длительной эволюции белков, но также и в том, что сравнительная численность разных ее видов изменяется во времени. Поэтому оценка соответствующих эффективных размеров популяций различных видов на основе наблюдаемых в настоящее время численностей может привести к серьезным ошибкам в определении размеров популяции, которые эти виды имели в прошлом.

Возможно, в данном случае больше подходит сравнение видов млекопитающих с совершенно разной продолжительностью поколения. Сравним, например, такие виды, как слон и мышь. Вероятно, скорости молекулярной эволюции (на год) у этих двух животных одинаковы. Примем, что продолжительность поколения у слона равна 20 годам, а у мыши — примерно 0,5 года; следовательно, средняя продолжительность поколения у слона в 40 раз выше, чем у мыши. Если к этим видам животных применима зависимость $g \propto 1/\sqrt{N_e}$, эффективный размер популяции мыши должен примерно в 1600 раз превышать эффективный размер популяции слона, что представляется нам

вполне реальным. В данной связи уместно рассмотреть следующий дополнительный вопрос: на какой стадии эволюционной истории этих видов животных возникла такая огромная разница продолжительностей их поколений? Другими словами, если предки слона и мыши в течение долгого времени после их дивергенции имели сходную продолжительность поколения, это необходимо принимать во внимание. Независимо от степени общности гипотезы постоянства скорости эволюции, если теория нейтральности верна, она должна соответствующим образом объяснять как случаи, когда гипотеза постоянства скорости эволюции справедлива, так и случаи (если таковые существуют), когда она не соответствует действительности.

Характерной особенностью молекулярной эволюции является также консервативная природа мутационных замещений: чем меньшими нарушениями существующей структуры и функций молекулы сопровождаются замещения, тем чаще они происходят в ходе эволюции. Наиболее отчетливо это было показано при исследовании эволюции белков, выявившем, что взаимные замены сходных по химическим свойствам аминокислот происходят чаще, чем несходных. С позиций теории нейтральной эволюции это легко объяснить, предположив, что вероятность того, что мутация будет нейтральной (т. е. не будет иметь неблагоприятного эффекта), выше, если она приводит к замещению сходной аминокислоты. Синонимичные замены, о которых я уже упоминал, — это предельный случай, когда различие между аминокислотами отсутствует.

С другой стороны, «селекционистское» объяснение, утверждающее, что мутационное изменение с меньшим фенотипическим проявлением имеет большую вероятность быть благоприятным (как это показано Фишером) и потому чаще «подхватывается» естественным отбором, совершенно неверно. Здесь упускается следующий ключевой момент: если эффект мутации очень мал, она будет иметь и очень небольшое селективное преимущество, а в соответствии с этим невысокую вероятность фиксации. На этом примере хорошо видно, что для осмысленного обсуждения механизмов молекулярной эволюции необходимо хорошо разбираться в стохастической теории популяционной генетики, в частности в теории, анализирующей поведение мутантных аллелей в конечной популяции: в любой популяции достаточно большого размера в каждом поколении возникает значительное число мутаций, однако большинство из них, включая мутации с небольшим селективным преимуществом, через несколько (порядка 10) поколений утрачиваются. Только небольшая доля мутаций распространяется по всей популяции и закрепляется в ней, причем этот процесс занимает очень долгое время (порядка 100 000 поколений). Во многих обсуждениях эволюции молчаливо предполагается, будто лю-

бая появляющаяся в популяции благоприятная мутация неизбежно закрепляется, однако это предположение неверно и не должно вводить нас в заблуждение при анализе механизмов молекулярной эволюции.

Рассмотрим теперь взаимосвязь между функциональными ограничениями и скоростью эволюции. Становится все более очевидно, что чем слабее функциональные ограничения, налагаемые на молекулу или ее часть, тем выше эволюционная скорость мутационных замещений. Грубо говоря, это означает, что молекулы или их части, менее важные в функциональном отношении, эволюционируют быстрее, чем более важные. Мы уже касались этого вопроса, когда обсуждали тот факт, что молекулярные изменения, которые с меньшей вероятностью испытывают давление отбора, в ходе эволюции возникают чаще. Примером такого рода служит средняя часть (пептид С) молекулы проинсулина, выщепляющаяся при образовании функционального инсулина, а также интроны в молекуле ДНК, транскрипты которых отсутствуют в окончательной структуре мРНК. Скорость эволюции таких участков, выраженная через число аминокислотных или нуклеотидных замен, очень высока. Мы также видели, что участки молекулы гемоглобина, расположенные на поверхности белковой глобулы, эволюционируют в 10 раз быстрее, чем расположенные в области «кармана» гема, в соответствии с тем известным фактом, что на наружную часть молекулы налагается меньше ограничений в смысле поддержания структуры и функции белка, чем на жизненно важный «карман».

Рассмотрим эту проблему количественно, основываясь на теории нейтральности. Предположим, что селективно нейтральные мутации составляют некоторую долю f_0 от всех мутаций, а остальные мутации ($1-f_0$) явно вредны. В редких случаях могут возникать и благоприятные мутации, однако мы считаем, что в определении общей скорости молекулярной эволюции ими можно пренебречь. Обозначим полную скорость мутирования на сайт через v_T ; тогда для нейтральных аллелей скорость мутирования будет равна $v = v_T f_0$. Из формулы $k = v$, применимой к нейтральным мутациям, мы имеем

$$k = v_T f_0. \quad (10.1)$$

Предположим теперь, что вероятность селективной нейтральности мутационного изменения сильно зависит от функциональных ограничений. Чем они слабее, тем выше вероятность f_0 нейтральности случайного изменения и тем больше k в уравнении (10.1). Очевидно, максимальная скорость эволюции достигается при $f_0 = 1$, т. е. в том случае, когда все мутации нейтральны ($k = v_T$). Здесь мы предполагаем, что суммарная скорость мутирования на сайт одинакова для разных молекул.

По моему мнению, очень высокая скорость эволюции, характерная для псевдоглобиновых генов млекопитающих (примерно $5 \cdot 10^{-9}$ на нуклеотидный сайт на год), должна быть очень близкой к этому пределу. Данная величина примерно в 10 раз выше, чем скорость аминокислотных замещений на сайт, приводящих к замещению аминокислот, в эволюции нормальных глобиновых генов. Кроме того, эта величина почти вдвое выше скорости синонимичных замен аминокислот, и, следовательно, синонимичные замены подвержены давлению отрицательного отбора (хотя интенсивность такого отбора должна быть крайне малой). Это согласуется с тем фактом, что синонимичные кодоны используются «неслучайным» (неравномерным) образом.

Неслучайное использование кодонов — это обычное явление; более того, существуют устойчивые типы их использования для различных генов в пределах генома данного вида. В настоящее время имеется много данных, свидетельствующих о том, что выбор синонимичного кодона в значительной степени зависит от наличия соответствующей тРНК. На основании такого рода данных была высказана гипотеза о соответствии характера предпочтительного использования кодонов содержанию в клетке соответствующих тРНК. Таким образом, концепция стабилизирующего отбора в количественной генетике может быть применима для рассмотрения данной проблемы с использованием метода диффузионных уравнений. Применение этого метода дает удовлетворительное количественное объяснение наблюдаемому характеру использования синонимичных кодонов и его влиянию на замедление фиксации синонимичных мутаций. Обнадешивает то, что такое универсальное явление, как неслучайное использование кодонов, получает удовлетворительное объяснение с позиций теории нейтральности с учетом селективных ограничений, тем более что данное явление часто приводится в качестве довода против этой теории.

Было показано, что если испытывающий давление стабилизирующего отбора количественный признак определяется большим числом расщепляющихся локусов (или сайтов), то средний коэффициент отбора на мутацию может быть крайне малым. Такие мутации имеют очень слабый неблагоприятный эффект, т. е. почти нейтральны, поэтому мутационные замещения в основном контролируются генетическим дрейфом, хотя в данном случае скорость эволюции может быть ниже, чем при полной нейтральности всех мутаций. Я считаю, что эти результаты свидетельствуют в пользу положения, состоящего в том, что, даже когда на фенотипическом уровне превалирует отбор, на молекулярном уровне может происходить интенсивная нейтральная эволюция. При обсуждении проблемы неслучайного использования синонимичных кодонов мы предполагали, что

тип использования кодонов в ходе эволюции изменяется таким образом, чтобы набор кодонов отвечал набору в клетке соответствующих тРНК. Возможно и обратное: относительные частоты разных тРНК (определяемые числом кодирующих их генов) в ходе эволюции изменяются таким образом, чтобы наблюдалось соответствие относительным частотам синонимичных кодонов в геноме вида. Какой из этих двух процессов является основным, неизвестно. Вполне вероятно, что они взаимосвязаны и очень медленно и согласованно эволюционируют. Это может служить хорошим примером «совместной эволюции» на молекулярном уровне — концепции, которая, как подчеркнул Комаи [281], важна для понимания механизма эволюции. Хотя Комаи в основном рассматривал роль совместной эволюции в ускорении эволюционных процессов на фенотипическом уровне, эта концепция применима также и к случаям, когда наблюдается необычайная эволюционная стабильность. Возможно, именно совместная эволюция ответственна за то, что определенный тип использования синонимичных кодонов бывает характерен для больших таксономических групп (например, для млекопитающих).

Эти соображения предполагают, что степень селективных ограничений в отношении синонимичных изменений по существу одинакова для разных генов одного генома. Отсюда становится ясно, почему скорости синонимичных замещений не только высоки, но и примерно одинаковы даже в случае генов, для продуктов которых характерны совершенно разные скорости замещения аминокислот. Например, α -тубулин позвоночных — это очень консервативный белок, скорость замещения для него равна $k_{aa} = 1,15 \cdot 10^{-11}$ на аминокислоту на год, что составляет лишь 1/100 от соответствующей скорости для гемоглобинов. Однако синонимичные компоненты скорости нуклеотидных замещений в третьем положении кодонов у α -тубулина и β -глобина практически совпадают. Такие же результаты были получены в отношении гистона H4 — наиболее консервативного из всех до сих пор изученных белков. Интересно, что синонимичные замены происходят у гистона H4 с очень высокой скоростью, сравнимой с соответствующей скоростью у других белков, скорости аминокислотных замещений у которых гораздо выше.

Существование максимальной скорости эволюции, определяемой суммарной скоростью мутирования в соответствии с формулой $k \leq v_t$, — одно из весьма удачных предсказаний теории нейтральности; его справедливость доказана результатами сравнительных исследований нуклеотидных последовательностей ДНК, которые в последние годы получили широкое развитие благодаря новым успехам молекулярной биологии. Кстати, основываясь на данных, полученных ранее методом гибридиза-

ции ДНК, сторонники селекционистской теории утверждали, что синонимичные замены встречаются реже, чем предсказывает теория нейтральности.

За последние несколько лет накопилось множество данных, свидетельствующих о наличии в геноме высших организмов большого количества дублированных (повторяющихся) последовательностей. Создается впечатление, что в геноме существуют дублированные копии для каждого гена. Эти факты подтверждают высказанное ранее предположение о том, что дубликация генов играет важную роль в прогрессивной эволюции: для формирования структур с более сложной организацией необходима дополнительная генетическая информация, а единственным способом получить ее является дубликация генов. Более того, существенно, что дубликации позволяют генам «проводить эволюционные эксперименты», которые без них были бы невозможны, так как благодаря наличию дубликатов гены освобождаются от постоянного давления естественного отбора, чье неуклонное влияние элиминирует варианты. Другими словами, в результате дубликации генов создаются условия, позволяющие генетическому дрейфу гораздо эффективнее действовать на мутации, а именно: многие мутации, явно вредные до дубликации, после нее становятся безвредными (т. е. селективно нейтральными) или очень слабо неблагоприятными, что способствует их распространению в популяции под действием генетического дрейфа. Затем некоторые из этих мутаций под действием изменившихся условий среды превращаются в полезные. Таким образом, теория нейтральности соприкасается и с проблемами прогрессивной эволюции. Процесс дубликации, облегчая образование новых генов, в то же время приводит к дегенерации одной из дублированных копий. Действительно, вероятность того, что дубликация гена приведет к его дегенерации, должна быть гораздо выше, чем вероятность того, что в результате дубликации образуется новый ген, обладающий полезной функцией.

В связи с дубликацией генов возникает важный вопрос: каким образом сама дубликация закрепляется в популяции? Широко распространенное мнение, что это связано с постоянным преимуществом гетерозигот, представляется ошибочным. Теория нейтральной эволюции дает гораздо более правдоподобное объяснение этому явлению. Сначала путем внутриврохромосомной дубликации создается тандемный повтор типа АА (где А — генный локус), который, будучи селективно нейтральным или очень слабо неблагоприятным, закрепляется в популяции в результате генетического дрейфа. Подобно обычным точковым мутациям генные дубликации все время возникают в популяции. Постоянный мутационный приток внутриврохромосомных дубликатов в сочетании с дрейфом генов в ходе длительной эволю-

ции, по-видимому, как раз и является основной причиной широкого распространения в геноме высших организмов повторяющихся последовательностей. Таким же образом можно объяснить наличие большого количества псевдогенов и другой не несущей никакой функции «ненужной» ДНК, а также эволюционное поведение «эгоистичной» ДНК (новые данные по этим вопросам см. в [89]).

Крайнюю форму генных дупликаций представляет собой семейство мультигенов, состоящее из большого числа тандемных повторов генов. Заслуживают внимания следующие семейства мультигенов: гены иммуноглобулина, гены рибосомной РНК и гистоновые гены. Для всех них характерна так называемая «согласованная», или «горизонтальная», эволюция. Это недавно обнаруженное явление состоит в том, что разные представители одного мультигенного семейства эволюционируют «в унисон», сохраняя внутривидовую гомогенность. Наиболее правдоподобное объяснение основано на предположении, что мутация, возникающая у одного из представителей семейства, случайно распространяется по всему семейству путем повторного внутрихромосомного неравного кроссинговера (горизонтальное распространение мутации). Этот процесс аналогичен случайной фиксации нейтральной мутации в популяции конечного размера в результате генетического дрейфа. Таким образом, эволюция семейства мультигенов может рассматриваться как двойной диффузионный процесс: распространение мутаций по хромосоме путем неравного кроссинговера или генной конверсии и распространение таких хромосом в популяции в результате обычного дрейфа генов, обусловленного выборкой гамет.

Обратимся теперь к проблеме внутривидовой генетической изменчивости. За последние 15 лет благодаря обширным электрофоретическим исследованиям полиморфизма белков было убедительно показано, что для природных популяций различных организмов характерен высокий уровень генетической изменчивости. Возникает вопрос: каков механизм поддержания такой высокой изменчивости? Многие популяционные генетики считают эту проблему одной из основных в своей области. Специалисты придерживаются по этому поводу совершенно разных мнений, однако их можно — с большей или меньшей степенью точности — отнести к двум противоположным лагерям: «селекционисты» утверждают, что большинство полиморфных аллелей активно поддерживается той или иной формой балансирующего отбора, тогда как «нейтралисты» считают, что такие аллели селективно нейтральны и «дрейфуют» в популяции случайным образом. Для проверки предполагаемой нейтральности полиморфных аллелей было разработано и поставлено множество экспериментов. Кроме того, был предложен целый ряд статистических тестов, с тем чтобы проверить соот-

ветствие наблюдаемых в природных популяциях распределений полиморфных аллелей предсказаниям теории нейтральности. Таким образом, теория нейтральности полиморфизма белков подверглась суровой и многосторонней проверке; кроме того, для объяснения полиморфизма были выдвинуты альтернативные гипотезы, основывающиеся на действии естественного отбора. В предыдущей главе я подробно рассмотрел, как с появлением новых данных постепенно разрешались кажущиеся противоречия теории нейтральности и накапливались факты, подтверждающие эту теорию.

Многие лабораторные данные, свидетельствующие о явных селективных различиях между полиморфными аллелями, по-видимому, ошибочны — нередко при более тщательной постановке эксперимента наличие отбора по ним не подтверждается. Кроме того, зачастую бывает неясно, действительно ли селективный эффект обусловлен полиморфными генами или же он связан с влиянием сцепленных с ними локусов. Утверждение, что адаптивную природу полиморфизма белков во всех случаях удастся неопровержимо доказать с помощью методов экологической генетики, оказалось необоснованным. До сих пор наиболее серьезные доказательства наличия естественного отбора получены в случае алкогольдегидрогеназы дрозофилы, однако результаты экспериментов противоречат друг другу, а последние тщательные исследования свидетельствуют о том, что если различия по приспособленности между аллелями *F* и *S* этого локуса и существуют, они, по-видимому, очень малы. Более того, наличие частотно-зависимого отбора также вызывает большие сомнения.

Для объяснения широко распространенного полиморфизма белков сторонники селекционистской теории все время привлекали гипотезу о сверхдоминировании, или преимуществе, гетерозигот. Однако за последние годы накопилось так много фактов против этой гипотезы (в том числе данные о широком распространении полиморфизма у гаплоидных видов), что ее уже нельзя считать научно обоснованной.

Для проверки гипотезы селективной нейтральности полиморфизма белков путем анализа распределений аллельных частот было разработано несколько статистических методов, однако оказалось, что каждый из них имеет свои недостатки. Никаких четких доказательств неадекватности теории нейтральности до сих пор получено не было. Напротив, проведенный Неем анализ взаимосвязи между средней и дисперсией гетерозиготности для различных организмов выявил полное соответствие с выводами теории нейтральности.

Одно время некоторые авторы утверждали, что неравновесие по сцеплению (т. е. неслучайное распределение аллелей между локусами) широко распространено в природных популя-

циях и что оно играет важную роль в поддержании полиморфизма белков. Однако по мере накопления экспериментальных данных становится все более очевидно, что между ферментными локусами у *Drosophila* практически не существует неравновесия по сцеплению, за исключением случаев ассоциации локусов с инверсиями. Более того, этот последний случай находит наиболее разумное объяснение в рамках концепции генетического дрейфа и влияния чисто случайных событий: можно показать математически, что, если скорость обмена аллелей между хромосомами с инверсией и хромосомами без инверсий значительно ниже, чем величина, обратная эффективному числу хромосом с инверсией, между этими аллелями и инверсиями возникает неслучайная ассоциация.

Есть основания полагать, что если в природе встречается сильное неравновесие по сцеплению, то локусы, по которым оно обнаруживается, сосредоточены в пределах очень небольшого участка хромосомы, где рекомбинация возникает исключительно редко. Кроме того, это неравновесие, очевидно, обусловлено не эпистатическим взаимодействием по приспособленности, а случайным дрейфом генов. Было высказано предположение, что неравновесие по сцеплению должно часто встречаться среди расщепляющихся нуклеотидных сайтов в пределах цистрона. Позднее Ота [540] показала, что между тесно сцепленными локусами в результате влияния генетического дрейфа возникает ярко выраженное неравновесие по сцеплению, если вид подразделен на демы, миграция между которыми ограничена. Ота предположила, что неравновесие по сцеплению, наблюдаемое в главном комплексе гистосовместимости человека и мыши, обусловлено именно этими причинами, а не эпистатическими взаимодействиями по приспособленности.

С точки зрения теории нейтральности внутривидовая генетическая изменчивость на молекулярном уровне, выражаемая в виде полиморфизма белков, является просто переходной фазой молекулярной эволюции. В таком случае этот полиморфизм должен контролироваться теми же механизмами, что и процесс молекулярной эволюции. Говоря более конкретно, уровень полиморфизма должен очень сильно зависеть от функциональных ограничений, налагаемых на данную молекулу, но почти не зависеть от условий среды, в которых обитает данный вид. Это утверждение находится в явном противоречии с одной из гипотез селекционизма, утверждающей, что полиморфизм представляет собой «адаптивную стратегию» вида, направленную на освоение гетерогенной среды обитания. Существуют разные варианты данной гипотезы, однако все они предсказывают наличие параллелизма между изменчивостью (гетерогенностью) окружающей среды и генетической изменчивостью, который, как оказалось, на самом деле отсутствует. В частно-

сти, положение, согласно которому организмы, обитающие на дне океанов, обладают крайне низкой степенью генетической изменчивости, поскольку их среда обитания весьма стабильна, не нашло подтверждения, причем наблюдаемые факты оказались прямо противоположными, т. е. генетическая изменчивость организмов, обитающих на дне океана, как правило, очень высока. После этого была предложена альтернативная гипотеза, согласно которой в условиях «стабильности трофических ресурсов» во временно устойчивых условиях среды происходит адаптация организмов к разнообразию пищевых ресурсов путем появления у них высокой степени генетической изменчивости. С другой стороны, высказывалось предположение, что гетерогенность среды благоприятна для аллелей, обладающих большой эволюционной пластичностью («гибкостью») и таким образом способствует мономорфизму. Подобные экологические гипотезы, предполагающие наличие «адаптивных стратегий», очень расплывчаты и с трудом поддаются математической интерпретации, поэтому я не останавливаюсь на них подробно.

Теория нейтральности предсказывает, что для локусов, кодирующих белки, на которые налагаются жесткие функциональные ограничения, характерен более низкий уровень генетической изменчивости. У человека и *Drosophila* мультимерные ферменты, образующие «межлокусные» гибридные молекулы, обладают значительно более низкой степенью полиморфизма, чем мономерные. Эти данные соответствуют положениям теории нейтральности, поскольку взаимодействие субъединиц таких мультимерных ферментов приводит к большим функциональным ограничениям и, следовательно, уменьшает вероятность того, что мутация будет безвредной, т. е. селективно нейтральной. С другой стороны, они опровергают предположение о том, что гибридные молекулы создают основу для сверхдоминирования. Это служит еще одним доказательством неадекватности гипотезы о сверхдоминировании как универсальной причине полиморфизма белков. У дрозофилы и других организмов для субстрат-специфичных ферментов в среднем характерен более низкий уровень полиморфизма, чем для субстрат-неспецифичных. Это вполне понятно, поскольку на субстрат-специфичные ферменты налагаются более жесткие функциональные ограничения. Недавно методом двумерного электрофореза было показано, что гетерозиготность по белкам, присутствующим в тканях в больших количествах, намного ниже, чем гетерозиготность по ферментам. Эти результаты можно объяснить тем, что мутационные изменения белков, присутствующих в клетке в больших количествах, влияют на приспособленность гораздо сильнее, чем изменения в белках, чье содержание в клетке мало. Действительно, к таким присутствующим

в больших количествах белкам относятся структурные белки — актин, тубулин и др., многие из которых, как известно, проявляют высокую эволюционную консервативность.

Если теория нейтральности справедлива, то в конечном счете мы увидим, что молекулы (или их части), важные в функциональном отношении и поэтому медленно эволюционирующие, имеют низкий уровень гетерозиготности. Это означает, что для некодирующих участков (особенно спейсеров), а также интронов (даже в случае генов, кодирующих белки) должна быть характерна высокая степень полиморфизма.

Теория нейтральности позволяет рассчитать степень гетерозиготности и полиморфизма с помощью простых формул. Рассмотрим модель диплоидной популяции со случайным скрещиванием, равным числом самцов и самок и постоянным во времени размером. Эффективная численность такой популяции N_e равна числу размножающихся особей в одном поколении. Рассмотрим какой-либо локус и предположим, что любая возникающая в этом локусе мутация ведет к появлению нового аллеля, не существовавшего в популяции ранее. Такая ситуация описывается в рамках «модели бесконечного числа аллелей». Введя эти упрощающие предположения, можно показать, что в условиях статистического равновесия, когда возникновение в результате мутаций новых аллелей уравнивается случайной элиминацией из популяции аллелей, существовавших в ней ранее, ожидаемая гетерозиготность равна

$$\bar{H}_e = \frac{4N_e v_0}{4N_e v_0 + 1}, \quad (10.2)$$

где v_0 — частота возникновения нейтральных мутаций на локус на поколение. При тех же предположениях вероятность того, что данный вид полиморфен по этому локусу, есть

$$P_{\text{поли}} = 1 - q^V, \quad (10.3)$$

где $V = 4N_e v_0$. В этой формуле q обозначает границу частоты, по которой определяется полиморфизм, т. е. популяция считается полиморфной по данному локусу, если частота наиболее распространенного аллеля этого локуса не превышает $1 - q$. Обычно q принимается равным 0,01; другими словами, популяция считается полиморфной, если суммарная частота менее распространенных аллелей превышает 1%.

Чтобы проиллюстрировать применение этих формул, представим себе гипотетическое млекопитающее с продолжительностью поколений, в точности равной 1 году; пусть численность вида этого млекопитающего составляет в каждом поколении 100 000 особей ($N_e = 10^5$). Рассмотрим генный локус, кодирующий некий условный белок, сходный по размерам с алькогольдегидрогеназой (т. е. состоящий из 333 аминокислот, от-

вечающих 999 нуклеотидным сайтам). Предположим далее, что суммарная скорость мутирования на нуклеотидный сайт на поколение (или на год) равна $5 \cdot 10^{-9}$ — величине, полученной при нашей оценке скорости эволюции псевдоглобиновых генов у млекопитающих. Известно, что в среднем около $1/3$ аминокислотных замещений в белке можно обнаружить, применяя стандартные электрофоретические методы. Тогда для нашего условного белка скорость возникновения электрофоретически выявляемых мутаций будет равна $1,3 \cdot 10^{-6}$. Значительная часть этих мутаций, вероятно, обладает неблагоприятным эффектом; поэтому предположим, что в среднем только $1/10$ мутаций, приводящих к изменению аминокислот, являются по существу нейтральными, тогда как остальные $9/10$ явно неблагоприятны и не вносят вклада ни в молекулярную эволюцию, ни в полиморфизм. Это предположение сделано в соответствии с тем фактом, что скорость эволюционных замещений аминокислот у нормальных глобинов равна примерно $1/10$ соответствующей величины для псевдоглобинов (отметим, что скорость эволюции глобинов является среди белков стандартной). Это согласуется также с соответствующей величиной, полученной при сравнении частот редких электрофоретических вариантов с частотами полиморфных вариантов в популяциях морской камбалы и японского макака. Тогда скорость мутирования для нейтральных аллелей на локус на поколение $v_0 = 1,3 \cdot 10^{-7}$, если учитывать только мутации, выявляемые методом электрофореза. При выбранной для нашего гипотетического млекопитающего численности популяции $N_e = 10^5$ имеем $4N_e v_0 = 5,2 \cdot 10^{-2}$. Используя приведенные выше формулы (10.2) и (10.3), получаем величину средней гетерозиготности $\bar{H}_e \approx 0,049$, или 4,9%, и долю полиморфных локусов $P \approx 0,213$, или 21,3%. Эти оценки хорошо согласуются с соответствующими величинами, полученными при исследовании большого числа видов животных (по оценке Нево $\bar{H}_e \approx 0,036 \pm 0,025$ и $P \approx 0,15 \pm 0,1$). Кроме того, можно построить теоретические распределения частот, которые также обнаруживают хорошее соответствие с наблюдаемыми распределениями.

Определение средней гетерозиготности очень важно с биологической точки зрения, поэтому соответствующие исследования проводились весьма интенсивно, и в настоящее время как для животных (позвоночных и беспозвоночных), так и для растений получено большое количество данных: наблюдаемые для природных популяций значения $\bar{H}_e$ лежат в пределах $0,0 \sim 0,3$ и редко превышают 30%. Одно время, когда гетерозиготность была определена для очень небольшого числа видов (человека, мыши, *Peromyscus*, мечехвоста и нескольких видов дрозофилы), теорию нейтральности часто подвергали критике на том осно-

вании, что средняя гетерозиготность оказывалась в интервале от 0,056 до 0,184, а это противоречило положениям теории нейтральности. Дело в том, что при этом получались значения $N_e v_0$, лежащие в интервале от 0,015 до 0,057, т. е. максимальное и минимальное значения $N_e v_0$ различались всего в четыре раза. Поскольку скорости мутационного процесса у различных организмов близки, это означает, что размеры популяций всех перечисленных выше организмов примерно одинаковы, что является совершенным абсурдом.

Теперь, когда получено гораздо больше данных, стало ясно, что приведенные доводы в основном потеряли свою убедительность. Прежде всего средняя гетерозиготность у некоторых организмов может быть очень низкой и даже равной нулю. В частности, низкий уровень гетерозиготности, вероятно, обычен для крупных млекопитающих. Одним из наиболее примечательных примеров является гренландский белый медведь, у которого все исследованные 13 локусов оказались мономорфными, т. е. $\bar{H}_e = 0$. С нашей точки зрения это легко понять, поскольку эффективная численность популяции белого медведя, по-видимому, очень мала. Действительно, при $N_e < 1000$ ожидаемая величина средней гетерозиготности становится менее 0,004.

Последний вопрос, на который еще нужно ответить в связи с упомянутой выше критикой, состоит в следующем: почему даже у видов с огромным размером популяций средняя гетерозиготность редко превышает 0,3. На этот вопрос можно ответить, заметив, что эффективная численность популяции, используемая при рассмотрении генетической изменчивости на молекулярном уровне, обычно гораздо меньше (иногда на несколько порядков) наблюдаемой. Дело в том, что эффективный размер популяции, определяемый за длительный период времени, может уменьшаться под действием целого ряда факторов: из-за неравенства числа размножающихся самцов и самок, колебаний численности популяции от поколения к поколению, различий в плодовитости, локального вымирания отдельных колоний с последующей реколонизацией и т. д. В частности, уровень средней гетерозиготности сильно падает, если размер популяции в ходе эволюции несколько раз резко уменьшается (эффект «горлышка бутылки») и если каждый раз после такого уменьшения численность популяции восстанавливается медленно. Даже если данный вид имеет в настоящее время весьма обширный ареал и состоит из огромного числа особей, маловероятно, чтобы в природных условиях такая же ситуация имела место в предшествовавшие миллионы лет и осталась неизменной в последующие миллионы лет. Рано или поздно она изменится в результате процесса видообразования. Таким образом, в общем случае величина N_e , используемая в формуле

(10.2) для оценки средней гетерозиготности, должна быть намного меньше N_e , определенной исходя из числа размножающихся в популяции особей в данный момент времени. Действительно, обширные исследования этого вопроса, проведенные Неем с сотрудниками, показывают, что всякий раз, когда удается грубо оценить эффективную численность популяции, наблюдаемая гетерозиготность оказывается ниже, чем это ожидается. Другими словами, теория нейтральности без труда объясняет «необыкновенно высокий» уровень генетической изменчивости.

Последние упомянутые результаты довольно трудно объяснить с селекционистских позиций, где предполагается, что аллельные варианты поддерживаются в популяциях благодаря позитивному влиянию факторов различной природы (включая частотно-зависимый отбор, действующий в пользу редких аллелей). Эти факторы, без сомнения, еще больше увеличивают степень гетерозиготности на фоне вклада мутационного процесса. С другой стороны, из теории нейтральности следует, что величины H_e и N_e должны сильно коррелировать, если исследовать зависимость между ними для большого числа видов.

В настоящее время в молекулярной биологии снова наблюдается некий бум, сопровождающийся быстрым накоплением новых данных. Это, без сомнения, приведет как к углублению, так и к расширению наших представлений о механизмах эволюции. Особого внимания заслуживает открытие интронно-экзонной структуры генов.

Говоря об этой структуре, Гилберт [128] предположил, что она может способствовать быстрым эволюционным изменениям, которые без нее были бы невозможными. Например, единственное мутационное изменение в сигнальном участке для сплайсинга может привести к образованию совершенно новых белков. Кроме того, Гилберт высказал предположение, что если экзоны соответствуют полипептидам, выполняющим разные функции, то рекомбинация внутри интронов приведет к независимой перегруппировке этих функций. Это предположение подтверждается результатами исследований Тонегавы и др. [500], согласно которым в гене константного участка тяжелой цепи иммуноглобулина имеются четыре отдельных экзона, кодирующие четыре разные единицы, три из которых имеют четко выраженную доменную структуру.

Ситуация с гемоглобинами не так ясна, как в случае иммуноглобинов, однако положение, которое было недавно выдвинуто Го [133] и полностью подтвердилось, поддерживает приведенную выше гипотезу. Го определял расстояние между атомами углерода в трехмерной структуре β -глобина мыши и идентифицировал четыре сегмента, обозначенные F1, F2, F3 и F4. Хорошо известно, что ген мыши, кодирующий β -глобин, состо-

ит из трех экзонов, разделенных двумя интронами, причем центральный сегмент (F2+F3) содержит группу гема. Основываясь на конформационных характеристиках, Го пришел к заключению, что участки F2 и F3 кодируются отдельными экзонами, и предположил, что предок теперешних генов α - и β -цепей глобинов содержал еще один интрон, который, возможно, был утрачен ввиду его бесполезности. Это положение получило убедительное подтверждение в работе [195], где показано, что в гене, кодирующем леггемоглобин у сон, имеются три интрона.

Интересно также отметить, что между членами одного семейства мультигенов может постоянно осуществляться (хотя и с низкой частотой) так называемая «перетасовка экзонов» („exon shuffling”), или «доменный перенос», на что указывают результаты сравнительного исследования константного участка тяжелой цепи гемоглобина [366]. Этот процесс может осуществляться путем двойного неравного кроссинговера или генной конверсии во время премейотических делений зародышевых клеток. Недавно Ота [451] предложила модель, учитывающую это явление, для объяснения необыкновенно высокой степени полиморфизма в главном комплексе гистосовместимости человека и мыши.

Я привел здесь лишь несколько примеров того, как в связи с новыми открытиями в области молекулярной биологии выявляются совершенно неожиданные аспекты эволюционных процессов. Они могут помочь нам проникнуть в глубь эволюционной истории, которую невозможно исследовать, пользуясь традиционными морфологическими и палеонтологическими методами.

Происхождение жизни на Земле было и остается одной из величайших научных загадок. До сих пор мы не можем однозначно ответить на вопрос, каким образом возникла первая самореплицирующаяся молекула (см., однако, работу [559], где представлены результаты важных экспериментов по репликации матрицы без участия ферментов в присутствии ионов металлов). С развитием молекулярной генетики в результате анализа новейших открытий на молекулярном уровне с эволюционистских позиций появилась надежда, что проблема происхождения жизни хоть немного прояснится. Начало этому было положено Эйгеном [93], который предложил и математически обосновал модель «гиперцикла» (четкое изложение теории Эйгена о происхождении генетической информации дано в работе [94]). Центральное положение в этой работе отводится концепции дарвиновского естественного отбора, однако не исключено, что важную роль здесь играют также стохастические процессы.

Возвращаясь к основной теме моей книги, отметим следующее. Математический аппарат, используемый мной для обоснования теории нейтральности, оперирует различными парамет-

рами: скоростью мутирования, коэффициентами отбора, численностью популяции, скоростью миграции и т. д. Хотя эти параметры нередко входят в произведения (например, $N_e v$), для сравнения теоретических выводов с данными экспериментов и наблюдений необходимо определить сами эти величины. Среди исследователей-эволюционистов распространено мнение, что точные значения упомянутых популяционно-генетических параметров определить невозможно, поэтому теории, в которых они используются, не имеют большой цены. Я придерживаюсь противоположного мнения: если мы действительно хотим понять механизм эволюции на молекулярном уровне, эти параметры необходимо исследовать и измерить. Ведь не утверждают же астрономы и космологи, что теории, оперирующие различными астрономическими параметрами, непригодны только потому, что эти параметры трудно точно измерить. Это напоминает мне афоризм, который, насколько я помню, принадлежит Галилею и в котором говорится: то, что мы можем измерить, следует измерить; то, что мы не можем измерить сейчас, следует стремиться сделать измеримым.

Многие прямо или косвенно говорили мне, что теория нейтральности не имеет биологического значения, поскольку нейтральные гены по определению не имеют отношения к адаптации. Роль нейтральных мутаций в эволюции с этой точки зрения часто описывается как «эволюционный шум». Мне кажется, такой подход слишком ограничен. Прежде всего самое важное в науке — поиски истины; если теория нейтральности не абсурдна как научная гипотеза, это уже является ценным. Далее, даже если «нейтральные аллели» функционально эквивалентны или почти эквивалентны в наиболее типичных условиях существования вида, вполне возможно, что при изменении условий среды на некоторые из этих аллелей начнет действовать отбор. Существуют эксперименты, убедительно подтверждающие такую возможность; я использовал для обозначения этого явления название «эффект Дикхьюзена — Хартла». Эффект Дикхьюзена — Хартла состоит в том, что нейтральные или почти нейтральные мутации могут иметь потенциальные возможности для отбора, который проявляется при соответствующих условиях. Таким образом, нейтральные мутации могут являться исходным материалом для адаптивной эволюции. В своей популярной книге об эволюции «Игра жизни» Колдер [36] пишет: «Ген, который сегодня является нейтральным, через тысячу поколений, в новых условиях среды, может стать очень важным». Он также замечает: «В течение многих миллионов лет в молекуле без вмешательства естественного отбора могут накапливаться незначительные изменения, однако настанет день, когда в полученной таким образом специфической молекуле произойдет еще одна мутация, которая придаст ей новое

значение — подобно тому, как прикуп одной карты может превратить не имеющий ценности набор карт в выигрышную комбинацию».

Материалом для адаптивной эволюции в основном являются мутантные гены, имеющие слабый неблагоприятный эффект, которые пассивно поддерживаются в генофонде вида и становятся благоприятными в связи с изменением внешней среды. Нейтральные мутации можно считать предельным случаем очень слабо неблагоприятных мутаций. В этой связи представляет интерес положение, выдвинутое Отой. Ота утверждает, что мутация может иногда (в строго определенных условиях среды) быть благоприятной; однако почти невероятно, что какая-либо мутация будет благоприятной в любых условиях среды. Таким образом, в постоянных условиях среды у мутации больше шансов оказаться благоприятной.

Хотя эта книга посвящена вопросам молекулярной эволюции, мне бы хотелось сделать несколько замечаний относительно эволюции на фенотипическом уровне. Я считаю, что дарвиновский (положительный) отбор действует в основном на фенотипы, которые являются конечным результатом действия многих генов. На этом уровне решающую роль в определении того, какие именно фенотипы будут отбираться, должны играть условия среды; положительному отбору безразлично, каков механизм связи данного фенотипа с определенным генотипом. Конечно, рассмотрение «экологических возможностей» необходимо для интерпретации данных по ископаемым; весьма маловероятно, что адаптивная фенотипическая эволюция происходит с равномерной скоростью. Я также хотел бы подчеркнуть, что законы, управляющие молекулярной эволюцией, совершенно отличаются от законов, управляющих эволюцией фенотипической.

Можно задать вопрос, почему на уровне фенотипов преобладает естественный отбор, тогда как на молекулярном уровне преимущественно происходит случайное закрепление нейтральных или почти нейтральных аллелей. Я думаю, что ответ на этот вопрос следует из того факта, что наиболее распространенным видом отбора на уровне фенотипов является стабилизирующий отбор. В отличие от того естественного отбора, который имел в виду Дарвин, стараясь объяснить эволюционные изменения, стабилизирующий отбор элиминирует особей с крайними фенотипами и оставляет особей, близких к популяционной средней. После ранних работ Бампуса на домовом воробье и Уэлдона на сухопутной улитке было обнаружено много случаев стабилизирующего отбора. Вероятно, лучшим примером подобного рода для человека является зависимость между весом новорожденных и их неонатальной смертностью, изученная Карн и Пенроузом. Как я уже упоминал, если в определении коли-

чественного признака участвует большое число расщепляющихся локусов или сайтов, средний коэффициент отбора на мутацию в условиях стабилизирующего отбора должен быть крайне малым. Было также показано, что такие мутации являются очень слабо вредными или почти нейтральными, так что мутационные замещения в основном контролируются генетическим дрейфом, хотя скорость «нейтральной эволюции» ниже, чем в случае, когда все мутации селективно нейтральны. Примененная к проблеме «неслучайного» или неравномерного использования синонимичных кодонов, эта модель предлагает удовлетворительное решение данной проблемы в рамках теории нейтральности.

Из всего сказанного выше вырисовывается следующая картина эволюции. Время от времени в результате изменения условий среды происходит сдвиг положения оптимума фенотипического признака; вид быстро отвечает на это, изменяя среднее значение данного признака. Во время этого короткого периода предполагается значительный сдвиг генных частот по многим локусам, но сам по себе данный процесс редко приводит к замещению генов. Однако большую часть времени наблюдается преобладание стабилизирующего отбора, при котором интенсивно происходит нейтральная эволюция, или случайная фиксация мутантных аллелей. Гоулд и Элдридж [141] утверждают, что эволюционная история представляет собой совокупность длительных периодов, характеризующихся относительной стабильностью и перемежающихся редкими периодами быстрых изменений («прерывистое равновесие»). Это находится в полном согласии с нарисованной мной выше картиной эволюции, где длительные периоды действия стабилизирующего отбора сменяются периодами интенсивного направленного отбора, следующего за изменением среды. Впрочем, как мне кажется, не обязательно привлекать для объяснения очень быстрых изменений другой механизм, как это делают Гоулд и Элдридж. Мы можем устранить противоречие между характером эволюции на фенотипическом уровне и на уровне молекулярном, предполагая преобладание стабилизирующего отбора на уровне фенотипов и участие очень большого числа нуклеотидных сайтов на молекулярном уровне.

В конце своей книги «Происхождение видов» Дарвин рассматривает густонаселенный берег реки со сложными экологическими взаимоотношениями между живыми организмами и отмечает, что все они обязаны своим появлением законам, «действующим вокруг нас», включая наследственность, изменчивость, борьбу за существование и естественный отбор. Особенно глубокое впечатление произвело на него образование удивительных высокоорганизованных форм в результате «борьбы природы с голодом и смертью». Дарвин заканчивает

свою книгу следующими хорошо известными словами:

«Есть величие в этом воззрении, по которому жизнь, с ее различными проявлениями, творец первоначально вдохнул в одну или ограниченное число форм; и между тем, как наша планета продолжает вращаться согласно неизменным законам тяготения, из такого простого начала развилось и продолжает развиваться бесконечное число самых прекрасных и самых изумительных форм».

Теперь мы знаем, что за этой удивительной картиной жизни скрыты огромные эволюционные изменения, которые происходили и происходят на генетическом уровне. Особенно интересным мне кажется то, что подавляющее большинство этих изменений обусловлено не естественным отбором, а случайным закреплением нейтральных или почти нейтральных мутаций. Хотя такие случайные процессы происходят слишком медленно и незаметно в рамках нашего кратковременного существования, в геологических масштабах времени их значение колоссально. Таким образом, время оставляет свои следы в геномах всех организмов на Земле. И это еще раз подтверждает величие наших воззрений на биологическую эволюцию.

Литература

1. Abramowitz M., Stegun I. A. (eds.). Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables, Washington, D. C., US Department of Commerce, 1964. [Имеется перевод: Абрамовиц М., Стиган И. Справочник по специальным функциям.— М.: Наука, 1979.]
2. Allendorf F. W. Protein polymorphism and the rate of loss of duplicate gene expression, *Nature, Lond.*, **272**, 76—78 (1978).
3. Allendorf F. W., Christiansen F. B., Dobson T., Eanes W. F., Frydenberg O. Electrophoretic variation in large mammals. I. The polar bear, *Thalarctos maritimus*, *Hereditas*, **91**, 19—22 (1979).
4. Allison A. C. Aspects of polymorphism in man, Cold Spring Harbor Symp. Quant. Biol., **20**, 239—255 (1955).
5. Alvarez L. W., Alvarez W., Asaro F., Michel H. V. Extraterrestrial cause for the Cretaceous-Tertiary extinction, *Science*, **208**, 1095—1108 (1980).
6. Ayala F. J. Darwinian versus non-Darwinian evolution in natural populations of *Drosophila*. In: Proceedings of the Sixth Berkeley Symp. Math. Stat. Prob., vol. 5, Le Cam L. M., Neyman J. and Scott E. L., eds., pp. 211—236, University of California Press, 1972.
7. Ayala F. J. Biological evolution: natural selection or random walk? *American Scientist*, **62**, 692—701 (1974).
8. Ayala F. J., Anderson W. W. Evidence of natural selection in molecular evolution, *Nature New Biol.*, **241**, 274—276 (1973).
9. Ayala F. J., Gilpin M. E. Lack of evidence for the neutral hypothesis of protein polymorphism, *J. Heredity*, **64**, 297—298 (1973).
10. Ayala F. J., Gilpin M. E. Gene frequency comparisons between taxa: support for the natural selection of protein polymorphisms, *Proc. Natl. Acad. Sci. USA*, **71**, 4847—4849 (1974).
11. Ayala F. J., Powell J. R., Tracey M. L., Mourão C. A., Pérez-Salas S. Enzyme variability in the *Drosophila willistoni* group. IV. Genetic variation in natural populations of *Drosophila willistoni*, *Genetics*, **70**, 113—139 (1972).
12. Ayala F. J., Tracey M. L. Genetic differentiation within and between species of the *Drosophila willistoni* group, *Proc. Natl. Acad. Sci. USA*, **71**, 999—1003 (1974).
13. Ayala F. J., Tracey M. L., Barr L. G., McDonald J. F., Pérez-Sala S. Genetic variation in natural populations of five *Drosophila* species and the hypothesis of the selective neutrality of protein polymorphisms, *Genetics*, **77**, 343—384 (1974).
14. Ayala F. J., Valentine J. W. Genetic variability in the pelagic environment: a paradox? *Ecology*, **60**, 24—29 (1979).
15. Ayala F. J., Valentine J. W., Hedgecock D., Barr L. G. Deep-sea asteroids: high genetic variability in a stable environment, *Evolution*, **29**, 203—212 (1975).
16. Baker M. C. Effective population size in a songbird: some possible implications, *Heredity*, **46**, 209—218 (1981).
17. Baltimore D. Gene conversion: some implications for immunoglobulin genes, *Cell*, **24**, 592—594 (1981).

18. Barnard E. A., Cohen M. S., Gold M. H., Kim J.-K. Evolution of ribonuclease in relation to polypeptide folding mechanism, *Nature*, Lond., **240**, 395—398 (1972).
19. Bell G. I., Pictet R. L., Rutter W. J., Cordell B., Tischer E., Goodman H. M. Sequence of the human insulin gene, *Nature*, Lond., **284**, 26—32 (1980).
20. Bird A. P. DNA methylation and the frequency of CpG in animal DNA, *Nucleic Acids Research*, **8**, 1499—1504 (1980).
21. Birky C. J., Skavaril R. V. Maintenance of genetic homogeneity in systems with multiple genomes, *Genet. Res. Camb.*, **27**, 249—265 (1976).
22. Black J. A., Gibson D. Neutral evolution and immunoglobulin diversity, *Nature*, Lond., **250**, 327—328 (1974).
23. Blundell T. L., Cutfield J. F., Cutfield S. M., Dodson E. J., Dodson G. G., Hodgkin D. C., Mercola D. A., Vijayan M. Atomic positions in rhombohedral 2-zinc insulin crystals, *Nature*, Lond., **231**, 506—511 (1971).
24. Blundell T. L., Wood S. P. Is the evolution of insulin Darwinian or due to selectively neutral mutation? *Nature*, Lond., **257**, 197—203 (1975).
25. Bogardt R. A., Jr., Jones B. N., Dwoilet F. E., Garner W. H., Lehman L. D., Gurd F. R. N. Evolution of the amino acid substitution in the mammalian myoglobin gene, *J. Mol. Evol.*, **15**, 197—218 (1980).
26. Bonnell M. L., Selander R. K. Elephant seals: genetic variation and near extinction, *Science*, **184**, 908—909 (1974).
27. Boyer S. H., Crosby E. F., Thrmon T. F., Noyes A. N., Fuller G. F., Leslie S. E., Shepard M. K., Herndon C. N. Hemoglobins A and A₂ in new world primates: comparative variation and its evolutionary implications, *Science*, **166**, 1428—1431 (1969).
28. Boyer S. H., Scott A. F., Kunkel L. M., Smith K. D. The proportion of all point mutations which are unacceptable: an estimate based on hemoglobin amino acid and nucleotide sequences, *Can. J. Genet. Cytol.*, **20**, 111—137 (1978).
29. Briscoe D. A., Robertson A., Malpica J.-M. Dominance at *Adh* locus in response of adult *Drosophila melanogaster* to environmental alcohol, *Nature*, Lond., **255**, 148—149 (1975).
30. Brown A. J. L., Langley C. H. Reevaluation of level of genic heterozygosity in natural population of *Drosophila melanogaster* by two-dimensional electrophoresis, *Proc. Natl. Acad. Sci. USA*, **76**, 2381—2384 (1979).
31. Brown A. H. D., Marshal D. R., Weir B. S. Current status of the charge state model for protein polymorphism. In: *Genetic Studies of Drosophila Populations*, Gibson J. B. and Oakeshott J. G., eds., pp. 15—43, Australian National University Press, 1981.
32. Bulmer M. G. The stability of equilibria under selection, *Heredity*, **27**, 157—162 (1971).
33. Bulmer M. G. Protein polymorphism, *Nature*, Lond., **234**, 410—411 (1971).
34. Bulmer M. G. The genetic variability of polygenic characters under optimizing selection, mutation and drift, *Genet. Res., Camb.*, **19**, 17—25 (1972).
35. Bulmer M. G. Geographical uniformity of protein polymorphisms, *Nature*, Lond., **241**, 199—200 (1973).
36. Calder N. *The Life Game*, London, British Broadcasting Corporation, 1973.
37. Calder N. *The Weather Machine*, London, British Broadcasting Corporation, 1974.
38. Cameron D. G., Vyse E. R. Heterozygosity in Yellowstone Park Elk, *Cervus canadensis*, *Biochemical Genetics*, **16**, 651—657 (1978).
39. Carlson E. A. *Genes, Radiation, and Society. The Life and Work of H. J. Muller*, Cornell University Press, 1981.
40. Cavalli-Sforza L. L. Population structure and human evolution, *Proc. Roy. Soc. London, Ser B*, **164**, 362—379 (1966).
41. Cavalli-Sforza L. L., Bodmer W. F. *The Genetics of Human Population*, San Francisco, W. H. Freeman and Co., 1971.

42. *Cavalli-Sforza L. L., Conterio F.* Analisi della fluttuazione di frequenze geniche nella popolazione della Val Parma, *Atti Assoc. Genet. Ital.*, **5**, 333—344 (1960).
43. *Cavalli-Sforza L. L., Edwards A. W. F.* Phylogenetic analysis: models and estimation procedures, *Amer. J. Hum. Genet.*, **19**, 233—257 (1967).
44. *Chakraborty R., Fuerst P. A., Nei M.* Statistical studies on protein polymorphism in natural populations II. Gene differentiation between populations, *Genetics*, **88**, 367—390 (1978).
45. *Charlesworth B.* Evidence against Fisher's theory of dominance, *Nature*, **278**, 848—849 (1979).
46. *Charlesworth B., Charlesworth D.* Dominance modification in a fluctuating environment, *Genet. Res., Camb.*, **34**, 281—285 (1979).
47. *Childs G., Maxson R., Cohn R. H., Kedes L.* Orphans: dispersed genetic elements derived from tandem repetitive genes of eucaryotes, *Cell*, **23**, 651—663 (1981).
48. *Christiansen F. B., Bundgaard J., Barker J. S. F.* On the structure of fitness estimates under post-observational selection, *Evolution*, **31**, 843—853 (1977).
49. *Clarke B.* Selective constraints on amino-acid substitutions during the evolution of proteins, *Nature*, **228**, 159—160 (1970).
50. *Clarke B.* Natural selection and the evolution of proteins, *Nature*, **232**, 487 (1971).
51. *Clarke B.* The contribution of ecological genetics to evolutionary theory: detecting the direct effects of natural selection on particular polymorphic loci, *Genetics*, **79**, 101—113 (1975).
52. *Clarke B.* The evolution of genetic diversity, *Proc. Roy. Soc. Lond., B*, **205**, 453—474 (1979).
53. *Clarke B., Allendorf F. W.* Frequency-dependent selection due to kinetic differences between allozymes, *Nature*, **279**, 732—734 (1979).
54. *Clarke B., O'Donald P.* Frequency-dependent selection, *Heredity*, **19**, 201—206 (1964).
55. *Cordell B., Bell G., Tischer E., DeNoto F. M., Ullrich A., Pictet R., Rutter W. J., Goodman H. M.* Isolation and characterization of a cloned rat insulin gene, *Cell*, **18**, 533—543 (1979).
56. *Coyne J.* Lack of genic similarity between two sibling species of *Drosophila* as revealed by varied techniques, *Genetics*, **84**, 593—607 (1976).
57. *Crick F.* Split genes and RNA slicing, *Science*, **204**, 264—271 (1979).
58. *Crow J. F.* Possible consequences of an increased mutation rate, *Eugenics Quarterly*, **4**, 67—80 (1957).
59. *Crow J. F.* Some possibilities for measuring selection intensities in man, *Human Biology*, **30**, 1—13 (1958).
60. *Crow J. F.* More on the heterozygous effects of lethals in *Drosophila* populations, *Amer. Natur.*, **98**, 447—449 (1964).
61. *Crow J. F.* The cost of evolution and genetic loads. In: *Haldane and Modern Biology*, Dronamraju K. R., ed., pp. 165—178, Baltimore, Johns Hopkins Press, 1968.
62. *Crow J. F.* Molecular genetics and population genetics, *Proc. XII International Congress of Genetics (Tokyo)*, vol. 3, pp. 105—113 (1969).
63. *Crow J. F.* Genetic loads and the cost of natural selection. In: *Mathematical Topics in Population Genetics*, Kojima K., ed., pp. 128—177, Berlin, Springer-Verlag, 1970.
64. *Crow J. F.* The dilemma of nearly neutral mutations: How important are they for evolution and human welfare? *J. Hered.*, **63**, 306—316 (1972).
65. *Crow J. F.* Minor viability mutants in *Drosophila*, *Genetics*, **92** (supplement), s165—s172 (1979).
66. *Crow J. F.* The neutralist-selectionist controversy: an overview. In: *Population and Biological Aspects of Human Mutation*, Hook E. B., ed., pp. 3—14, New York, Academic Press, 1981.

67. Crow J. F., Kimura M. An Introduction to Population Genetics Theory, New York, Harper and Row, 1970.
68. Crow J. F., Kimura M. The effective number of a population with overlapping generations: A correction and further discussion, *Amer. J. Hum. Genet.*, **24**, 1—10 (1972).
69. Crow J. F., Kimura M. Efficiency of truncation selection, *Proc. Natl. Acad. Sci. USA*, **76**, 396—399 (1979).
70. Crow J. F., Temin R. G. Evidence for the partial dominance of recessive lethal genes in natural populations of *Drosophila*, *Amer. Natur.*, **98**, 21—33 (1964).
71. Daly K., Clarke B. Selection associated with the alcohol dehydrogenase locus in *Drosophila melanogaster*: differential survival of adults maintained on low concentrations of ethanol, *Heredity*, **46**, 219—226 (1981).
72. Darwin C. The Origin of Species by Means of Natural Selection, London, John Murray, 1859.
73. Dawkins R. The Selfish Gene, Oxford University Press, 1976.
74. Dayhoff M. O. (ed.). Atlas of Protein Sequence and Structure, vol. 4, Silver Spring, Maryland, National Biomedical Research Foundation, 1969.
75. Dayhoff M. O. (ed.). Atlas of Protein Sequence and Structure, vol. 5, Washington, DC., National Biomedical Research Foundation, 1972.
76. Dayhoff M. O. (ed.). Atlas of Protein Sequence and Structure, vol. 5, supplement 2, Washington, DC., National Biomedical Research Foundation, 1976.
77. Dayhoff M. O. (ed.). Atlas of Protein Sequence and Structure, vol. 5, supplement 3, Washington, DC., National Biomedical Research Foundation, 1978.
78. Dayhoff M. O., Eck R. V. Atlas of Protein Sequence and Structure, 1967—1968, Silver Spring, National Biomedical Research Foundation, 1968.
79. Deevey E. S., Jr. The human population, *Scientific American*, **203**(3), 195—204 (1960).
80. Dickerson R. E. The structure of cytochrome *c* and the rates of molecular evolution, *J. Mol. Evol.*, **1**, 26—45 (1971).
81. Dickerson R. E., Geis I. The Structure and Action of Proteins, Menlo Park, W. A. Benjamin, 1969.
82. Dobzhansky Th. Genetics and the Origin of Species, 1st edn., Columbia University Press, 1937.
83. Dobzhansky Th. Genetics and the Origin of Species, 3rd edn., Columbia University Press, 1951.
84. Dobzhansky Th. A review of some fundamental concepts and problems of population genetics, Cold Spring Harbor Symp. Quant. Biol., **20**, 1—15 (1955).
85. Dobzhansky Th. Genetic diversity and fitness. In: Genetics Today (Proc. XI International Congr. Genetics), Geerts S. J., ed., pp. 541—552, Oxford, Pergamon Press, 1965.
86. Dobzhansky Th. Genetics of the Evolutionary Process, Columbia University Press, 1970.
87. Dolan R., Robertson A. The effect of conditioning the medium in *Drosophila* in relation to frequency dependent selection, *Heredity*, **35**, 311—316 (1975).
88. Doolittle W. F., Sapienza C. Selfish genes, the phenotype paradigm and genome evolution, *Nature, Lond.*, **284**, 601—603 (1980).
89. Dover G. A., Flavell R. B. (eds.). Genome Evolution, London, Academic Press, 1982.
90. Dover G. Molecular drive: a cohesive mode of species evolution, *Nature*, **299**, 111—117 (1982).
91. Dykhuizen D., Hartl D. L. Selective neutrality of 6PGD allozymes in *E. coli* and the effects of genetic background, *Genetics*, **96**, 801—817 (1980).
92. Efstratiadis A., Kafatos F. C., Maniatis T. The primary structure of rabbit β -globin mRNA as determined from cloned DNA, *Cell*, **10**, 571—585 (1977).

93. *Eigen M.* Selforganization of matter and the evolution of biological macromolecules, *Naturwissenschaften*, **58**, 465—523 (1971).
94. *Eigen M., Gardiner W., Schuster P., Winkler-Oswatitsh R.* The origin of genetic information, *Scientific American*, **244**(2), 78—94 (1981).
95. *Epstein C. J.* Non-randomness of amino-acid changes in the evolution of homologous proteins, *Nature, Lond.*, **215**, 355—359 (1967).
96. *Ewens W. J.* Population Genetics, London, Methuen, 1969.
97. *Ewens W. J.* The sampling theory of selectively neutral alleles, *Theor. Popul. Biol.*, **3**, 87—112 (1972).
98. *Ewens W. J.* Population genetics theory in relation to the neutralist-selectionist controversy, *Adv. Hum. Genet.* **8**, 67—134 (1977).
99. *Ewens W. J.* Mathematical Population Genetics, Berlin, Heidelberg, Springer-Verlag, 1979.
100. *Ewens W. J., Feldman M. W.* Analysis of neutrality in protein polymorphism, *Science*, **183**, 446—448 (1974).
101. *Felsenstein J.* On the biological significance of the cost of gene substitution, *Amer. Natur.*, **105**, 1—11 (1971).
102. *Ferris S. D., Whitt G. S.* Loss of duplicate gene expression after polyploidization, *Nature, Lond.*, **265**, 258—260 (1977).
103. *Fincham J. R. S.* Gene Complementation, Amsterdam, Benjamin, 1966.
104. *Fisher R. A.* The correlation between relatives on the supposition of Mendelian inheritance, *Trans. Roy. Soc. Edinburgh*, **52**, 399—433 (1918).
105. *Fisher R. A.* On the dominance ratio, *Proc. Roy. Soc. Edinburgh*, **42**, 321—341 (1922).
106. *Fisher R. A.* The distribution of gene ratios for rare mutations, *Proc. Roy. Soc. Edinburgh*, **50**, 205—220 (1930).
107. *Fisher R. A.* The Genetical Theory of Natural Selection, Oxford University Press, 2nd revised edition (1958), Dover Publications, 1930.
108. *Fisher R. A.* The measurement of selective intensity, *Proc. Roy. Soc. London, Ser. B*, **121**, 58—62 (1936).
109. *Fisher R. A.* Population genetics, *Proc. Roy. Soc. London, Ser. B*, **141**, 510—523 (1953).
110. *Fisher R. A., Ford E. B.* The Sewall Wright effect, *Heredity*, **4**, 117—119 (1950).
111. *Fisher W. K., Nash A. R., Thompson E. O. P.* Haemoglobins of the shark. *Heterodontus portusjacksoni* III. Amino acid sequence of the β -chain, *Aust. J. Biol. Sci.*, **30**, 487—506 (1977).
112. *Fitch W. M.* Evolutionary variability in hemoglobins. In: *Synthese, Struktur und Funktion des Hämoglobins*, Martin H. and Nowicki L., eds., pp. 199—215, München, J. F. Lehmanns Verlag, 1971.
113. *Fitch W. M.* Rate of change of concomitantly variable codons, *J. Mol. Evol.*, **1**, 84—96 (1971).
114. *Fitch W. M.* Toward defining the course of evolution: minimum change for a specific tree topology, *Syst. Zool.*, **20**, 406—416 (1971).
115. *Fitch W. M.* Estimating the total number of nucleotide substitutions since the common ancestor of a pair of homologous genes: comparison of several methods and three beta hemoglobin messenger RNA's, *J. Mol. Ecol.*, **16**, 153—209 (1980).
116. *Fitch W. M., Langley C. H.* Protein evolution and the molecular clock, *Fed. Proc.*, **35**, 2092—2097 (1976).
117. *Fitch W. M., Markowitz E.* An improved method for determining codon variability in a gene and its application to the rate of fixation of mutations in evolution, *Biochem. Genet.*, **4**, 579—593 (1970).
118. *Fontdevila A., Mendez J., Ayala F. J., McDonald J.* Maintenance of allozyme polymorphisms in experimental populations of *Drosophila*, *Nature, Lond.*, **255**, 149—151 (1975).
119. *Ford E. B.* Ecological Genetics, 1st edn., London, Methuen, 2nd and 3rd edns. (1965, 1971), London, Chapman and Hall, 1964.
120. *Ford E. B.* Genetic Polymorphism, London, Faber and Faber, 1965.

121. Franklin I., Lewontin R. C. Is the gene the unit of selection? *Genetics*, **65**, 707—734 (1970).
122. Fuerst P. A., Chakraborty R., Nei M. Statistical studies on protein polymorphism in natural populations I. Distribution of single locus heterozygosity, *Genetics*, **86**, 455—483 (1977).
123. Fuerst P. A., Ferrell R. E. The stepwise mutation model: an experimental evaluation utilizing hemoglobin variants, *Genetics*, **94**, 185—201 (1980).
124. Gaines M. S., Whittam T. Genetic changes in fluctuating vole populations: selective vs. nonselective forces, *Genetics*, **96**, 767—778 (1980).
125. Gattin L. L. Counter-examples to a neutralist hypothesis, *J. Mol. Evol.*, **7**, 185—195 (1976).
126. Gauss D. H., Sprinzl M. Compilation of tRNA sequences, *Nucl. Acids Res.*, **9**, r1—r23 (1981).
127. Gibson J. B., Lewis N., Adena M. A., Wilson S. R. Selection for ethanol tolerance in two populations of *Drosophila melanogaster* segregating alcohol dehydrogenase allozymes, *Aust. J. Biol. Sci.*, **32**, 387—398 (1979).
128. Gilbert W. Why genes in pieces? *Nature*, **271**, 501 (1978).
129. Gillespie J. H. Sampling theory for alleles in a random environment, *Nature*, **266**, 443—445 (1977).
130. Gillespie J. H., Kojima K. The degree of polymorphisms in enzymes involved in energy production compared to that in nonspecific enzymes in two *Drosophila ananassae* populations, *Proc. Natl. Acad. Sci. USA*, **61**, 582—585 (1968).
131. Gillespie J. H., Langley C. H. A general model to account for enzyme variation in natural populations, *Genetics*, **76**, 837—848 (1974).
132. Ginzburg I., Behar L., Givol D., Littauer U. Z. The nucleotide sequence of rat α -tubulin: 3'-end characteristics, and evolutionary conservation, *Nucl. Acids Res.*, **9**, 2691—2697 (1981).
133. Gō M. Correlation of DNA exonic regions with protein structure units in haemoglobin, *Nature*, **291**, 90—92 (1981).
134. Gojobori T. Studies of protein polymorphism by data analyses with respect to loci. Doctorial Dissertation, Faculty of Science, Kyushu University, 1979.
135. Gojobori T. Means and variances of heterozygosity and protein function. In: *Molecular Evolution, Protein Polymorphism and the Neutral Theory*, Kimura M., ed., pp. 137—148, Tokyo, Japan Scientific Societies Press Berlin, Springer-Verlag, 1982.
136. Gooch J. L., Schopf T. J. M. Genetic variability in the deep sea: relation to environmental variability, *Evolution*, **26**, 545—552 (1972).
137. Goodman M. Substitutional trends and non-random changes in rates during protein evolution. In: *Evolution of Protein Molecules*, Matsubara H. and Yamanaka T., eds., pp. 17—32, Tokyo, Japan Scientific Societies Press, 1978.
138. Goodman M., Moore G. W., Barnabas J., Matsuda G. The phylogeny of human globin genes investigated by the maximum parsimony method, *J. Mol. Evol.*, **3**, 1—48 (1974).
139. Goodman M., Moore G. W., Matsuda G. Darwinian evolution in the genealogy of hemoglobin, *Nature*, **253**, 603—608 (1975).
140. Goossens M., Dozy A. M., Embury S. H., Zachariades Z., Hadjiminis M., Stamatoyannopoulos G., Kan Y. W. Triplicated α -globin loci in humans, *Proc. Natl. Acad. Sci. USA*, **77**, 518—521 (1980).
141. Gould S. J., Eldredge N. Punctuated equilibria: the tempo and mode of evolution reconsidered, *Paleobiology*, **3**, 115—151 (1977).
142. Grant V. *Organismic Evolution*, San Francisco, W. H. Freeman, 1977. [Имеется перевод: Грант В. Эволюция организмов.— М.: Мир, 1980.]
143. Grantham R. Amino acid difference formula to help explain protein evolution, *Science*, **185**, 862—864 (1974).
144. Grantham R. Workings of the genetic code, *Trends in Biochemical Sciences*, **5**, 327—331 (1980).

145. Grantham R., Gautier C., Gouy M. Codon frequencies in 119 individual genes confirm consistent choices of degenerate bases according to genome type, Nucl. Acids Res., 8, 1893—1912 (1980).
146. Grantham R., Gautier C., Gouy M., Jacobzone M., Mercier R. Codon catalog usage is a genome strategy modulated for gene expressivity, Nucl. Acids Res., 9, r43—r74 (1981).
147. Grantham R., Gautier C., Gouy M., Mercier R., Pavé A. Codon catalog usage and the genome hypothesis, Nucl. Acids Res., 8, r49—r62 (1980).
148. Grunstein M., Grunstein J. E. The histone H4 gene of *Strongylocentrotus purpuratus*: DNA and mRNA sequences at the 5' end, Cold Spring Harbor Symp. Quant. Biol., 42 (part 2), 1083—1092 (1978).
149. Grunstein M., Schedl P., Kedes L. Sequence analysis and evolution of sea urchin (*Lytechinus pictus* and *Strongylocentrotus purpuratus*) histone H4 messenger RNAs, J. Mol. Biol., 104, 351—369 (1976).
150. Guess H. A., Ewens W. J. Theoretical and simulation results relating to the neutral allele theory, Theor. Popul. Biol., 3, 434—437 (1972).
151. Haga T. Structure and dynamics of natural populations of a diploid *Trillium*. In: Chromosomes Today, Darlington C. D. and Lewis K. R., eds., pp. 207—217, Edinburgh, Oliver and Boyd, 1969.
152. Haga T., Kurabayashi M. Genome and polyploidy in the genus *Trillium*. V. Chromosomal variation in natural populations of *Trillium kamtschaticum* Pall, Mem. Fac. Sci. Kyushu Univ., Ser. E, Biol. 1, 159—185 (1954).
153. Haldane J. B. S. A mathematical theory of natural and artificial selection. Part I. Trans. Camb. Phil. Soc., 23, 19—41 (1924).
154. Haldane J. B. S. A mathematical theory of natural and artificial selection. Part V. Selection and mutation, Proc. Camb. Phil. Soc., 23, 838—844 (1927).
155. Haldane J. B. S. A note on Fisher's theory of the origin of dominance, Amer. Natur., 64, 87—90 (1930).
156. Haldane J. B. S. A mathematical theory of natural and artificial selection. Part VIII. Metastable populations, Proc. Camb. Phil. Soc., 27, 137—142 (1931).
157. Haldane J. B. S. The Causes of Evolution, New York, Harper and Row, 1932. [Имеется перевод: Холден Дж. Факторы эволюции. М.—Л: Биомедгиз, 1935.]
158. Haldane J. B. S. The effect of variation on fitness, Amer. Natur., 71, 337—349 (1937).
159. Haldane J. B. S. Forty years of genetics. In: Background to Modern Science, Needham J. and Pagel W., eds., pp. 225—243, Cambridge University Press, 1938.
160. Haldane J. B. S. Suggestions as to quantitative measurement of rates of evolution, Evolution, 3, 51—56 (1949).
161. Haldane J. B. S. The statics of evolution. In: Evolution as a Process, Huxley J. et al., eds., pp. 109—121, London, George Allen and Unwin, 1954.
162. Haldane J. B. S. Karl Pearson, 1857—1957, Biometrika, 44, 303—313 (1957).
163. Haldane J. B. S. The cost of natural selection, J. Genet., 55, 511—524 (1957).
164. Haldane J. B. S. The Unity and Diversity of Life. (Sardar Vallabhbhai Patel Lectures) Delhi, Ministry of Information and Broadcasting, Government of India, 1958.
165. Haldane J. B. S. Natural selection. In: Darwin's Biological Work, Bell P. R., ed., pp. 101—149, Cambridge University Press, 1959.
166. Haldane J. B. S. A defense of beanbag genetics, Perspect. Biol. Med., 7, 343—359 (1964).
167. Haldane J. B. S., Jayakar S. D. Polymorphism due to selection of varying direction, J. Genet., 58, 237—242 (1963).
168. Hardy G. H. Mendelian proportions in a mixed population, Science, 28, 49—50 (1908).

169. Harris H. Enzyme polymorphism in man, Proc. Roy. Soc. London, Ser. B, **164**, 298—310 (1966).
170. Harris H. Molecular evolution: the neutralist-selectionist controversy, Fed. Proc., **35**, 2079—2082 (1976).
171. Harris H., Hopkinson D. A. Average heterozygosity per locus in man: an estimate based on the incidence of enzyme polymorphisms, Ann. Hum. Genet., **36**, 9—20 (1972).
172. Harris H., Hopkinson D. A., Edwards Y. H. Polymorphism and subunit structure of enzymes: a contribution to the neutralist-selectionist controversy, Proc. Natl. Acad. Sci. USA, **74**, 698—701 (1977).
173. Harris H., Hopkinson D. A., Robson E. B. The incidence of rare alleles determining electrophoretic variants: data on 43 enzyme loci in man, Ann. Hum. Genet., **37**, 237—253 (1974).
174. Hartl D. L., Dykhuizen D. E. Potential for selection among nearly neutral allozymes of 6-phosphogluconate dehydrogenase in *Escherichia coli*, Proc. Natl. Acad. Sci. USA, **78**, 6344—6348 (1981).
175. Hayashi A., Stamatoyannopoulos G. Role of penultimate tyrosine in haemoglobin β subunit, Nature, New Biol., **235**, 70—72 (1972).
176. Hedrick P. W., Ginevan M. F., Ewing E. P. Genetic polymorphism in heterogeneous environments, Ann. Rev. Ecol. Syst., **7**, 1—32 (1976).
177. Heindell H. C., Liu A., Paddock G. V., Studnicka G. M., Salser W. A. The primary sequence of rabbit α -globin mRNA, Cell, **15**, 43—54 (1978).
178. Hendrickson W. A., Love W. E. Structure of lamprey haemoglobin, Nature, New Biol., **232**, 197—203 (1971).
179. Hildreth P. E. The problem of synthetic lethals in *Drosophila melanogaster*, Genetics, **41**, 729—742 (1956).
180. Hill W. G. Effective size of populations with overlapping generations, Theor. Pop. Biol., **3**, 278—289 (1972).
- 180a. Hill W. G. A note on effective population size with overlapping generations, Genetics, **92**, 317—322 (1979).
181. Ho M. W., Saunders P. T. Beyond neo-Darwinism — an epigenetic approach to evolution, J. Theor. Biol., **78**, 573—591 (1979).
182. Holmquist R. The method of parsimony: an experimental test and theoretical analysis of the adequacy of molecular restoration studies, J. Mol. Biol., **135**, 939—958 (1979).
183. Holmquist R., Jukes T. H., Pangburns S. Evolution of transfer RNA, J. Mol. Biol., **78**, 91—116 (1973).
184. Hood L., Campbell J. H., Elgin S. C. R. The organization, expression, and evolution of antibody genes and other multigene families, Annu. Rev. Genet., **9**, 305—353 (1975).
185. Hori H., Osawa S. Evolutionary change in 5S RNA secondary structure and a phylogenetic tree of 54 5S RNA species, Proc. Natl. Acad. Sci. USA, **76**, 381—385 (1979).
186. Horuk R., Goodwin P., O'Connor K., Neville R. W. J., Lazarus N. R., Stone D. Evolutionary change in the insulin receptors of hystricomorph rodents, Nature, Lond., **279**, 439—440 (1979).
187. Huang S. L., Singh M., Kojima K. A study of frequency-dependent selection observed in the esterase-6 locus of *Drosophila melanogaster* using a conditioned media method, Genetics, **68**, 97—104 (1971).
188. Hunt L. T., Hurst-Calderone S., Dayhoff M. O. Globins. In: Atlas of Protein Sequence and Structure, vol. 5, supplement 3, Dayhoff M. O., ed., Washington, D. C., National Biomedical Research Foundation, 1978.
189. Ikemura T. The frequency of codon usage in *E. coli* genes: correlation with abundance of cognate tRNA. In: Genetics and Evolution of RNA Polymerase, tRNA and Ribosomes, Osawa S., Ozeki H., Uchida H. and Yura T., eds., pp. 519—523, University of Tokyo Press, 1980.
190. Ikemura T. Correlation between the abundance of *Escherichia coli* transfer RNAs and the occurrence of the respective codons in its protein genes, J. Mol. Biol., **146**, 1—21 (1981).

191. Ikemura T. Correlation between the abundance of *Escherichia coli* transfer RNAs and the occurrence of the respective codons in its protein genes: a proposal for a synonymous codon choice that is optimal for the *E. coli* translational system, *J. Mol. Biol.*, **151**, 389—409 (1981).
192. Ingram V. M. *The Hemoglobins in Genetics and Evolution*, Columbia University Press, 1963.
193. Iserberg I. Histones, *Annu. Rev. Biochem.*, **48**, 159—191 (1979).
194. Jahn C. L., Hutchison C. A., III, Phillips S. J., Weaver S., Haigwood N. L., Volvia C. F., Edgell M. H. DNA sequence organization of the β -globin complex in the BALB/c mouse, *Cell*, **21**, 159—168 (1980).
195. Jensen E. O., Paudan K., Hyldig-Nielsen J. J., Jørgensen P., Marcker K. A. The structure of a chromosomal leghaemoglobin gene from soybean, *Nature*, Lond., **291**, 677—679 (1981).
196. Johnson G. B. Evidence that enzyme polymorphisms are not selectively neutral, *Nature*, New Biol., **237**, 170—171 (1972).
197. Johnson G. B. On the estimation of effective number of alleles from electrophoretic data, *Genetics*, **78**, 771—776 (1974).
198. Jones J. S. How much genetic variation? *Nature*, Lond., **288**, 10—11 (1980).
199. Jukes T. H. Recent advances in studies of evolutionary relationships between proteins and nucleic acids, *Space Life Sci.*, **1**, 469—490 (1969).
200. Jukes T. H. Comparison of the polypeptide chains of globins, *J. Mol. Evol.*, **1**, 46—62 (1971).
201. Jukes T. H. Codons and nearest-neighbor nucleotide pairs in mammalian messenger RNA, *J. Mol. Evol.*, **11**, 121—127 (1978).
202. Jukes T. H. Neutral changes during divergent evolution of hemoglobins, *J. Mol. Evol.*, **11**, 267—269 (1978).
203. Jukes T. H. Dr. Best, insulin, and molecular evolution, *Can. J. Biochem.*, **57**, 455—458 (1979).
204. Jukes T. H. Silent nucleotide substitutions and the molecular evolutionary clock, *Science*, **210**, 973—978 (1980).
205. Jukes T. H. Silent nucleotide substitutions during evolution, *Naturwissenschaften*, **67**, 534—539 (1980).
206. Jukes T. H., Cantor C. R. Evolution of protein molecules. In: *Mammalian Protein Metabolism III*, Munro H. N., ed., pp. 21—132, New York, Academic Press, 1969.
207. Jukes T. H., Holmquist R. Evolutionary clock: nonconstancy of rate in different species, *Science*, **177**, 530—532 (1972).
208. Jukes T. H., King J. L. Deleterious mutations and neutral substitutions, *Nature*, Lond., **231**, 114—115 (1971).
209. Jukes T. H., King J. L. Evolutionary loss of ascorbic acid synthesizing ability, *J. Hum. Evol.*, **4**, 85—88 (1975).
210. Jukes T. H., King J. L. Evolutionary nucleotide replacement in DNA, *Nature*, Lond., **281**, 605—606 (1979).
211. Karlin S., Lieberman U. Random temporal variation in selection intensities: case of large population size, *Theor. Pop. Biol.*, **6**, 355—382 (1974).
212. Karn M. N., Penrose L. S. Birth weight and gestation time in relation to maternal age, parity and infant survival, *Ann. Eugen.*, **16**, 147—164 (1951).
213. Kedes L. H. Histone genes and histone messengers, *Annu. Rev. Biochem.*, **48**, 837—870 (1979).
214. Kerr R. A. Asteroid theory of extinctions strengthened, *Science*, **210**, 514—517 (1980).
215. Kerr W. E. Multiple alleles and genetic load in bees, *J. Apicult. Res.*, **6**, 61—64 (1967).
216. Kesten H. The number of alleles in electrophoretic experiments, *Theor. Pop. Biol.*, **18**, 290—294 (1980).
217. Kettlewell H. B. D. Selection experiments on industrial melanism in the *Lepidoptera*, *Heredity*, **9**, 323—342 (1955).

218. Kettlewell H. B. D. Further selection experiments on industrial melanism in the *Lepidoptera*, *Heredity*, **10**, 287—301 (1956).
219. Kettlewell H. B. D. A survey of the frequencies of *Biston betularia* (L.) (Lep.) and its melanic forms in Great Britain, *Heredity*, **12**, 51—72 (1958).
220. Kim S. H., Quigley G. J., Suddath F. L., McPherson A., Sneden D., Kim J. J., Weinzierl J., Rich A. Three-dimensional structure of yeast phenylalanine transfer RNA: folding of the polynucleotide chain, *Science*, **179**, 285—288 (1973).
221. Kim S. H., Suddath F. L., Quigley G. J., McPherson A., Sussman J. L., Wang A. H. J., Seeman N. C., Rich A. Three-dimensional tertiary structure of yeast phenylalanine transfer RNA, *Science*, **185**, 435—440 (1974).
222. Kimura M. Solution of a process of random genetic drift with a continuous model, *Proc. Natl. Acad. Sci. USA*, **41**, 144—150 (1955).
223. Kimura M. Some problems of stochastic processes in genetics, *Ann. Math. Stat.*, **28**, 882—901 (1957).
224. Kimura M. Optimum mutation rate and degree of dominance as determined by the principle of minimum genetic load, *J. Genet.*, **57**, 21—34 (1960).
225. Kimura M. Genetic load of a population and its significance in evolution (Japanese with English summary), *Jap. J. Genet.*, **35**, 7—33 (1960).
226. Kimura M. Some calculations on the mutational load, *Jap. J. Genet.*, **36**, (supplement), 179—190 (1961).
227. Kimura M. On the probability of fixation of mutant genes in a population, *Genetics*, **47**, 713—719 (1962).
228. Kimura M. Diffusion models in population genetics, *J. Appl. Probab.*, **1**, 177—232 (1964).
229. Kimura M. Evolutionary rate at the molecular level, *Nature, Lond.*, **217**, 624—626 (1968).
230. Kimura M. Genetic variability maintained in a finite population due to mutational production of neutral and nearly neutral isoalleles, *Genet. Res., Camb.*, **11**, 247—269 (1968).
231. Kimura M. The rate of molecular evolution considered from the standpoint of population genetics, *Proc. Nat. Acad. Sci. USA*, **63**, 1181—1188 (1969).
232. Kimura M. The number of heterozygous nucleotide sites maintained in a finite population due to steady flux of mutations, *Genetics*, **61**, 893—903 (1969).
233. Kimura M. The length of time required for a selectively neutral mutant to reach fixation through random frequency drift in a finite population, *Genet. Res., Camb.*, **15**, 131—133 (1970).
234. Kimura M. Theoretical foundation of population genetics at the molecular level, *Theor. Pop. Biol.*, **2**, 174—208 (1971).
235. Kimura M. Gene pool of higher organisms as a product of evolution, *Cold Spring Harbor Symp. Quant. Biol.*, **38**, 515—524 (1974).
236. Kimura M. Preponderance of synonymous changes as evidence for the neutral theory of molecular evolution, *Nature, Lond.*, **267**, 275—276 (1977).
237. Kimura M. Change of gene frequencies by natural selection under population number regulation, *Proc. Natl. Acad. Sci. USA*, **75**, 1934—1937. (1978).
238. Kimura M. Model of effectively neutral mutations in which selective constraint is incorporated, *Proc. Natl. Acad. Sci. USA*, **76**, 3440—3444 (1979).
239. Kimura M. The neutral theory of molecular evolution, *Scientific American*, **241**(5), 94—104 (1979).
240. Kimura M. Average time until fixation of a mutant allele in a finite population under continued mutation pressure: studies by analytical, numerical, and pseudosampling methods, *Proc. Natl. Acad. Sci. USA*, **77**, 522—526 (1980).
241. Kimura M. A simple method for estimating evolutionary rates of base

- substitutions through comparative studies of nucleotide sequences, *J. Mol. Evol.*, **16**, 111—120 (1980).
242. Kimura M. Was globin evolution very rapid in its early stages?: a dubious case against the rate-constancy hypothesis, *J. Mol. Evol.*, **17**, 110—113 (1981).
 243. Kimura M. Doubt about studies of globin evolution based on maximum parsimony codons and the augmentation procedure, *J. Mol. Evol.*, **17**, 121—122 (1981).
 244. Kimura M. Estimation of evolutionary distances between homologous nucleotide sequences, *Proc. Natl. Acad. Sci. USA*, **78**, 454—458 (1981).
 245. Kimura M. Possibility of extensive neutral evolution under stabilizing selection with special reference to non-random usage of synonymous codons, *Proc. Natl. Acad. Sci. USA*, **78**, 5773—5777 (1981).
 246. Kimura M., Crow J. F. The number of alleles that can be maintained in a finite population, *Genetics*, **49**, 725—738 (1964).
 247. Kimura M., Crow J. F. Effect of overall phenotypic selection on genetic change at individual loci, *Proc. Natl. Acad. Sci. USA*, **75**, 6168—6171 (1978).
 248. Kimura M., King J. L. Fixation of a deleterious allele at one of two «duplicate» loci by mutation pressure and random drift, *Proc. Natl. Acad. Sci. USA*, **76**, 2858—2861 (1979).
 249. Kimura M., Maruyama T. The mutational load with epistatic gene interactions in fitness, *Genetics*, **54**, 1337—1351 (1966).
 250. Kimura M., Maruyama T. The substitutional load in a finite population, *Heredity*, **24**, 101—114 (1969).
 251. Kimura M., Maruyama T. Pattern of neutral polymorphism in a geographically structured population, *Genet. Res., Camb.*, **18**, 125—131 (1971).
 252. Kimura M., Maruyama T., Crow J. F. The mutational load in small populations, *Genetics*, **48**, 1303—1312 (1963).
 253. Kimura M., Ohta T. The average number of generations until fixation of a mutant gene in a finite population, *Genetics*, **61**, 763—771 (1969).
 254. Kimura M., Ohta T. The average number of generations until extinction of an individual mutant gene in a finite population, *Genetics*, **63**, 701—709 (1969).
 255. Kimura M., Ohta T. Genetic load at a polymorphic locus which is maintained by frequency-dependent selection, *Genet. Res., Camb.*, **16**, 145—150 (1970).
 256. Kimura M., Ohta T. Protein polymorphism as a phase of molecular evolution, *Nature, Lond.*, **229**, 467—469 (1971).
 257. Kimura M., Ohta T. *Theoretical Aspects of Population Genetics*, Princeton, Princeton University Press, 1971.
 258. Kimura M., Ohta T. On the stochastic model for estimation of mutational distance between homologous proteins, *J. Mol. Evol.*, **2**, 87—90 (1972).
 259. Kimura M., Ohta T. Mutation and evolution at the molecular level, *Genetics*, **73** (supplement), 19—35 (1973).
 260. Kimura M., Ohta T. The age of a neutral mutant persisting in a finite population, *Genetics*, **75**, 199—212 (1973).
 261. Kimura M., Ohta T. On some principles governing molecular evolution, *Proc. Natl. Acad. Sci. USA*, **71**, 2848—2852 (1974).
 262. Kimura M., Ohta T. Distribution of allelic frequencies in a finite population under stepwise production of neutral alleles, *Proc. Natl. Acad. Sci. USA*, **72**, 2761—2764 (1975).
 263. Kimura M., Ohta T. Further comments on «Counter-examples to a neutralist hypothesis», *J. Mol. Evol.*, **9**, 367—368 (1977).
 264. Kimura M., Ohta T. Stepwise mutation model and distribution of allelic frequencies in a finite population, *Proc. Natl. Acad. Sci. USA*, **75**, 2868—2872 (1978).
 265. Kimura M., Ohta T. Population genetics of multigene family with special reference to decrease of genetic correlation with distance between gene

- members on a chromosome, Proc. Natl. Acad. Sci. USA, 76, 4001—4005 (1979).
266. Kimura M., Weiss G. M. The stepping stone model of population structure and the decrease of genetic correlation with distance, Genetics, 49, 561—576 (1964).
 267. King J. L. Continuously distributed factors affecting fitness, Genetics, 55, 483—492 (1967).
 268. King J. L., Jukes T. H. Non-Darwinian evolution, Science, 164, 788—798 (1969).
 269. King J. L., Ohta T. Polyallelic mutational equilibria, Genetics, 79, 681—691 (1975).
 270. King M. C., Wilson A. C. Evolution at two levels: Molecular similarities and biological differences between human and chimpanzees, Science, 188, 107—116 (1975).
 271. Kingman J. F. C. Coherent random walks arising in some genetical models, Proc. Roy. Soc. London, Ser. A, 351, 19—31 (1976).
 272. Kingman J. F. C. A note on multidimensional models of neutral mutation, Theor. Pop. Biol., 11, 285—290 (1977).
 273. Klose J., Feller M. Genetic variability of proteins from plasma membranes and cytosols of mouse organs, Biochem. Genet., 19, 859—870 (1981).
 274. Klug A., Ladner J., Robertus J. D. The structural geometry of co-ordinated base changes in transfer RNA, J. Mol. Biol., 89, 511—516 (1974).
 275. Koch R. E. The influence of neighboring base pairs upon base-pair substitution mutation rates, Proc. Natl. Acad. Sci. USA, 68, 773—776 (1971).
 276. Koehn R. K., Eanes W. F. Subunit size and genetic variation of enzymes in natural populations of *Drosophila*, Theor. Pop. Biol., 11, 330—341 (1977).
 277. Koehn R. K., Eanes W. F. Molecular structure and protein variation within and among populations. In: Evolutionary Biology, vol. 11, Hecht M. K., Steere W. C. and Wallace B., eds., pp. 39—100.
 278. Kojima K., Gillespie J. H., Tobari Y. N. A profile of *Drosophila* species' enzymes assayed by electrophoresis. 1. Number of alleles, heterozygosities, and linkage disequilibrium in glucose metabolizing systems and some other enzymes, Biochem. Genet., 4, 627—637 (1970).
 279. Kojima K., Tobari Y. N. The pattern of viability changes associated with genotype frequency at the alcohol dehydrogenase locus in a population of *Drosophila melanogaster*, Genetics, 61, 201—209 (1969).
 280. Kojima K., Yarbrough K. M. Frequency-dependent selection at the esterase 6 locus in *Drosophila melanogaster*, Proc. Natl. Acad. Sci. USA, 57, 645—649 (1967).
 281. Komai T. Theory of «mutual selection», Proc. Japan Acad. 46, 383—386 (1970).
 282. Kondo S. Evolutionary considerations on DNA repair and mutagenesis. In: Molecular Evolution and Polymorphism, Kimura M., ed., pp. 313—331, Mishima, National Institute of Genetics, 1977.
 283. Konkel D. A., Tilghman S. M., Leder P. The sequence of the chromosomal mouse β -globin major gene: homologs in capping, splicing and poly (A) sites, Cell, 15, 1125—1132 (1978).
 284. Kortlandt A. New Perspectives on Ape and Human Evolution, University of Amsterdam, 1972.
 285. Krimbas C. B., Tsakas S. The genetics of *Dacus oleae*. V. Changes of esterase polymorphism in a natural population following insecticide control — selection or drift? Evolution, 25, 454—460 (1971).
 286. Kurtén B. The Age of Mammals, London, Weidenfeld and Nicolson, 1971.
 287. Lacy E., Maniatis T. The nucleotide sequence of a rabbit β -globin pseudogene, Cell, 21, 545—553 (1980).
 288. Laki K. Enzymatic effects of thrombin, Fed. Proc., 24, 794—799 (1965).
 289. Lande R. Effective deme sizes during long-term evolution estimated from rates of chromosomal rearrangement, Evolution, 33, 234—251 (1979).

290. Lande R. Genetic variation and phenotypic evolution during allopatric speciation, *Amer. Natur.*, **116**, 463—479 (1980).
291. Langley C. H., Fitch W. M. An examination of the constancy of the rate of molecular evolution, *J. Mol. Ecol.*, **3**, 161—177 (1974).
292. Langley C. H., Smith D. B., Johnson F. M. Analysis of linkage disequilibria between allozyme loci in natural populations of *Drosophila melanogaster*, *Genet. Res., Camb.*, **32**, 215—229 (1978).
293. Langridge J. Mutation spectra and the neutrality of mutations, *Aust. J. Biol. Sci.*, **27**, 309—319 (1974).
294. Latter B. D. H. Influence of selection pressures on enzyme polymorphisms in *Drosophila*, *Nature, Lond.*, **257**, 590—591 (1975).
295. Latter B. D. H. Gene frequency distributions in species of *Drosophila*. In: Genetic Studies of *Drosophila* Populations, Gibson J. B. and Oakesbott J. G., eds., pp. 3—13, Australian National University Press, 1981.
296. Latter B. D. H. The distribution of heterozygosity in temperate and tropical species of *Drosophila*, *Genet. Res., Camb.*, **38**, 137—156 (1981).
297. Lawn R. M., Efstratiadis A., O'Connell C., Maniatis T. The nucleotide sequence of the human β -globin gene, *Cell*, **21**, 647—651 (1980).
298. Leder P., Hansen J. N., Konkel D., Leder A., Nishioka Y., Talkington C. Mouse globin system: a functional and evolutionary analysis, *Science*, **209**, 1336—1342 (1980).
299. Lerner I. M. Genetic Homeostasis, Edinburgh, Oliver and Boyd, 1954.
300. Lesk A. M., Chothia C. How different amino acid sequences determine similar protein structures: the structure and evolutionary dynamics of the globins, *J. Mol. Biol.*, **136**, 225—270 (1980).
301. Levins R. Evolution in Changing Environments, Princeton University Press, 1968.
302. Lewontin R. C. The Genetic Basis of Evolutionary Change, Columbia University Press, 1974. [Имеется перевод: Левонтин Р. Генетические основы. — М.: Мир, 1978.]
303. Lewontin R. C., Ginzburg L. R., Tuljapurkar S. D. Heterosis as an explanation for large amounts of genic polymorphism, *Genetics*, **88**, 149—170 (1978).
304. Lewontin R. C., Hubby J. L. A molecular approach to the study of genic heterozygosity in natural populations, II. Amount of variation and degree of heterozygosity in natural populations of *Drosophila pseudoobscura*, *Genetics*, **54**, 595—609 (1966).
305. Lewontin R. C., Krakauer J. Distribution of gene frequency as a test of the theory of the selective neutrality of polymorphisms, *Genetics*, **74**, 175—195 (1973).
306. Lewontin R. C., Krakauer J. Testing the heterogeneity of F values, *Genetics*, **80**, 397—398 (1975).
307. Li W.-H. A mixed model of mutation for electrophoretic identity of proteins within and between populations, *Genetics*, **83**, 423—432 (1976).
308. Li W.-H. Stochastic Models in Population Genetics, Benchmark Papers in Genetics Series, vol. 7, Stroudsburg, Pennsylvania, Dowden, Hutchinson and Ross, 1977.
309. Li W.-H. Distribution of nucleotide differences between two randomly chosen cistrons in a finite population, *Genetics*, **85**, 331—337 (1977).
310. Li W.-H. Maintenance of genetic variability under the joint effect of mutation, selection and random drift, *Genetics*, **90**, 349—382 (1978).
311. Li W.-H. Maintenance of genetic variability under the pressure of neutral and deleterious mutations in a finite population, *Genetics*, **92**, 647—667 (1979).
312. Li W.-H., Gojobori T., Nei M. Pseudogenes as a paradigm of neutral evolution, *Nature, Lond.*, **292**, 237—239 (1981).
313. Li W.-H., Nei M. Drift variances of heterozygosity and genetic distance in transient states, *Genet. Res., Camb.*, **25**, 229—248 (1975).
314. Lomedico P., Rosenthal N., Efstratiadis A., Gilbert W., Kolodner R., Ti-

- zard R. The structure and evolution of the two nonallelic rate preproinsulin genes, *Cell*, **18**, 545—558 (1979).
315. Loukas M., Krimbas C. B., Vergini Y. The genetics of *Drosophila subobscura* populations. IX. Studies on linkage disequilibrium in four natural populations, *Genetics*, **93**, 497—523 (1979).
 316. Loukas M., Vergini Y., Krimbas C. B. The genetics of *Drosophila subobscura* populations, XVII. Further genic heterogeneity within electromorphs by urea denaturation and the effect of the increased genic variability on linkage disequilibrium studies, *Genetics*, **97**, 429—441 (1981).
 317. Lueders K., Leder A., Leder P., Kuff E. Association between a transposed α -globin pseudogene and retrovirus-like elements in the BALB/c mouse genome, *Nature, Lond.*, **295**, 426—428 (1982).
 318. MacArthur R. H., Wilson E. O. The Theory of Island Biogeography, Princeton University Press, 1967.
 319. McConkey E. H., Taylor B. J., Phan D. Human heterozygosity: a new estimate, *Proc. Natl. Acad. Sci. USA*, **76**, 6500—6504 (1979).
 320. McDonald J. F., Ayala F. J. Genetic response to environmental heterogeneity, *Nature, Lond.*, **250**, 572—574 (1974).
 321. McFarland W. N., Pough F. H., Cade T. J., Heiser J. B. Vertebrate Life, New York, Macmillan, London, Collier Macmillan, 1979.
 322. McKenzie J. A., McKechnie S. W. Ethanol tolerance and the *Adh* polymorphism in a natural population of *Drosophila melanogaster*, *Nature, Lond.*, **272**, 75—76 (1978).
 323. McLachlan A. D. Repeating sequences and gene duplication in proteins, *J. Mol. Biol.*, **64**, 417—437 (1972).
 324. McLaughlin P. J., Dayhoff M. O. Evolution of species and proteins: a time scale. In: Atlas of Protein Sequence and Structure 1972, Dayhoff M. O., ed., pp. 47—52, Washington, D. C., National Biomedical Research Foundation, 1972.
 325. Maita T., Tanaka E., Goodman M., Matsuda G. Amino acid sequences of the α and β chains of adult hemoglobin of the tupai, *Tupai agilis*, *J. Biochem.*, **82**, 603—605 (1977).
 326. Malécot G. Identical loci and relationship. In: Proc. Fifth Berkeley Symp. Math. Sta. Prob. vol. 4, pp. 317—332, Berkeley, University of California Press, 1967.
 327. Margoliash E., Smith E. L. Structure and functional aspects of cytochrome *c* in relation to evolution. In: Evolving Genes and Proteins, Bryson V. and Vogel H. J., eds., pp. 221—242, New York, Academic Press, 1965.
 328. Marotta C. A., Wilson J. T., Forget B. G., Weissman S. M. Human β -globin messenger RNA. III. Nucleotide sequences derived from complementary DNA, *J. Biol. Chem.*, **252**, 5040—5053 (1977).
 329. Marshall D. R., Brown A. H. D. The charge-state model of protein polymorphisms in natural populations, *J. Mol. Evol.*, **6**, 149—163 (1975).
 330. Martial J. A., Hallewell R. A., Baxter J. D., Goodman H. M. Human growth hormone: Complementary DNA cloning and expression in bacteria, *Science*, **205**, 602—607 (1979).
 331. Maruyama T. On the fixation probability of mutant genes in a subdivided population, *Genet. Res., Camb.*, **15**, 221—226 (1970).
 332. Maruyama T. Effective number of alleles in a subdivided population, *Theor. Pop. Biol.*, **1**, 273—306 (1970).
 333. Maruyama T. An invariant property of a structured population, *Genet. Res., Camb.*, **18**, 81—84 (1971).
 334. Maruyama T. Some invariant properties of a geographically structured finite population: Distribution of heterozygotes under irreversible mutation, *Genet. Res., Camb.*, **20**, 141—149 (1972).
 335. Maruyama T. Rate of decrease of genetic variability in a two-dimensional continuous population of finite size, *Genetics*, **70**, 639—651 (1972).
 336. Maruyama T. The age of an allele in a finite population, *Genet. Res., Camb.*, **23**, 137—143 (1974).

337. *Maruyama T.* Stochastic Problems in Population Genetics (Lecture Notes in Biomathematics 17), Berlin, Springer-Verlag, 1977.
338. *Maruyama T., Kimura M.* Some methods for treating continuous stochastic processes in population genetics, *Jap. J. Genet.*, **46**, 407—410 (1971).
339. *Maruyama T., Kimura M.* Geographical uniformity of selectively neutral polymorphisms, *Nature, Lond.*, **249**, 30—32 (1974).
340. *Maruyama T., Kimura M.* Moments for sum of an arbitrary function of gene frequency along a stochastic path of gene frequency change, *Proc. Natl. Acad. Sci. USA*, **72**, 1602—1604 (1975).
341. *Maruyama T., Kimura M.* Theoretical study of genetic variability, assuming stepwise production of neutral and very slightly deleterious mutations, *Proc. Natl. Acad. Sci. USA*, **75**, 919—922 (1978).
342. *Maruyama T., Kimura M.* Genetic variability and effective population size when local extinction and recolonization of subpopulations are frequent, *Proc. Natl. Acad. Sci. USA*, **77**, 6710—6714 (1980).
343. *Mather K.* The genetical structure of populations, *Symp. Soc. Exp. Biol.*, **7**, 66—95 (1953).
344. *Mather K.* Genetic Structure of Populations, London, Chapman and Hall, 1973.
345. *Matsubara H., Hase T., Wakabayashi S., Wada K.* Gene duplications during evolution of chloroplast-type ferredoxins. In: *Evolution of Protein Molecules*, Matsubara H. and Yamanaka T., eds., pp. 209—220, Tokyo, Japan Scientific Soc. Press, Center for Academic Publications, 1978.
346. *Matsuda G.* Molecular evolution of hemoglobins. In: *Evolution of Protein Molecules*, Matsubara H. and Yamanaka T., eds., pp. 175—185, Tokyo, Japan Scientific Soc. Press, 1978.
347. *Matsuda H., Gojobori T., Takahata N.* Theoretical study on protein polymorphism and its bearing on the evolution of protein molecules. In: *Evolution of Protein Molecules*, Matsubara H. and Yamanaka T., eds., pp. 89—100, Tokyo, Japan Scientific Soc. Press, 1978.
348. *Maynard Smith J.* «Haldane's dilemma» and the rate of evolution, *Nature, Lond.*, **219**, 1114—1116 (1968).
349. *Maynard Smith J.* Protein polymorphism, *Nature, New Biol.*, **237**, 31 (1972).
350. *Maynard Smith J., Hoekstra R.* Polymorphism in a varied environment: how robust are the models? *Genet. Res., Camb.*, **35**, 45—57 (1980).
351. *Mayr E.* Integration of genotypes: synthesis, *Cold Spring Harbor Symp. Quant. Biol.*, **20**, 327—333 (1955).
352. *Mayr E.* Where are we? *Cold Spring Harbor Symp. Quant. Biol.*, **24**, 1—14 (1959).
353. *Mayr E.* Animal Species and Evolution, Harvard University Press, 1963. [Имеется перевод: Майр Э. Зоологический вид и эволюция. — М.: Мир, 1968.]
354. *Mayo O.* The role of duplications in evolution, *Heredity*, **25**, 543—553 (1970).
355. *Maxam A. M., Gilbert W.* A new method for sequencing DNA, *Proc. Natl. Acad. Sci. USA*, **74**, 560—564 (1977).
356. *Milkman R.* Heterosis as a major cause of heterozygosity in nature, *Genetics*, **55**, 493—495 (1967).
357. *Milkman R.* Electrophoretic variation in *Escherichia coli* from natural sources, *Science*, **182**, 1024—1026 (1973).
358. *Milkman R.* Selection is the major determinant, *Trend Biochem. Sci.*, **1**, N152—N154 (1976).
359. *Milkman R.* Selection differentials and selection coefficients, *Genetics*, **88**, 391—403 (1978).
360. *Minawa A., Birley A. J.* The genetical response to natural selection by varied environments. 1. Short-term observations, *Heredity*, **40**, 39—50 (1978).

361. Miyata T., Hayashida H. Extraordinarily high evolutionary rate of pseudogenes: evidence for the presence of selective pressure against changes between synonymous codons, *Proc. Natl. Acad. Sci. USA*, **78**, 5739—5743 (1981).
362. Miyata T., Miyazawa S., Yasunaga T. Two types of amino acid substitution in protein evolution, *J. Mol. Evol.*, **12**, 219—236 (1979).
363. Miyata T., Yasunaga T. Molecular evolution of mRNA: a method for estimating evolutionary rates of synonymous and amino acid substitutions from homologous nucleotide sequences and its application, *J. Mol. Evol.*, **16**, 23—36 (1980).
364. Miyata T., Yasunaga T. Rapidly evolving mouse α -globin-related pseudogene and its evolutionary history, *Proc. Natl. Acad. Sci. USA*, **78**, 450—453 (1981).
365. Miyata T., Yasunaga T., Nishida T. Nucleotide sequence divergence and functional constraint in mRNA evolution, *Proc. Natl. Acad. Sci. USA*, **77**, 7328—7332 (1980).
366. Miyata T., Yasunaga T., Yamawaki-Kataoka Y., Obata M., Honjo T. Nucleotide sequence divergence of mouse immunoglobulin γ_1 and γ_2 chain genes and the hypothesis of intervening sequence-mediated domain transfer, *Proc. Natl. Acad. Sci. USA*, **77**, 2143—2147 (1980).
367. Modiano G., Battistuzzi G., Motulsky A. G. Nonrandom patterns of codon usage and of nucleotide substitutions in human α - and β -globin genes: an evolutionary strategy reducing the rate of mutations with drastic effects? *Proc. Natl. Acad. Sci. USA*, **78**, 1110—1114 (1981).
368. Moran P. A. P. Wandering distributions and the electrophoretic profile, *Theor. Pop. Biol.*, **8**, 318—330 (1975).
369. Moran P. A. P. Wandering distributions and the electrophoretic profile. II. *Theor. Pop. Biol.*, **10**, 145—149 (1976).
370. Morgan P. Frequency-dependent selection at two enzyme loci in *Drosophila melanogaster*, *Nature, Lond.*, **263**, 765—766 (1976).
371. Mross G. A., Doolittle R. F. Amino acid sequence studies on artiodactyl fibrinopeptides II. Vicuna, elk, muntjak, pronghorn antelope, and water buffalo, *Arch. Biochem. Biophys.*, **122**, 674—684 (1967).
372. Mukai T. The genetic structure of natural populations of *Drosophila melanogaster*. I. Spontaneous mutation rate of polygenes controlling viability, *Genetics*, **50**, 1—19 (1964).
373. Mukai T. Lack of experimental evidence supporting selection for the maintenance of isozyme polymorphism in *Drosophila melanogaster*. In: *Molecular Evolution and Polymorphism*, Kimura M., ed., pp. 103—126, Mishima, National Institute of Genetics, 1977.
374. Mukai T., Chigusa S. I., Mettler L. E., Grow J. F. Mutation rate and dominance of gene affecting viability in *Drosophila melanogaster*, *Genetics*, **72**, 335—355 (1972).
375. Mukai T., Mettler L. E., Chigusa S. I. Linkage disequilibrium in a local population of *Drosophila melanogaster*, *Proc. Natl. Acad. Sci. USA*, **68**, 1065—1069 (1971).
376. Mukai T., Schaffer H. E., Cockerham C. C. Genetic consequences of truncation selection at the phenotypic level in *Drosophila melanogaster*, *Genetics*, **72**, 763—769 (1972).
377. Mukai T., Tachida H., Ichinose M. Selection for viability at loci controlling protein polymorphisms in *Drosophila melanogaster* is very weak at most, *Proc. Natl. Acad. Sci. USA*, **77**, 4857—4860 (1980).
378. Mukai T., Voelker R. A. The genetic structure of natural populations of *Drosophila melanogaster*. XIII. Further studies on linkage disequilibrium, *Genetics*, **86**, 175—185 (1977).
379. Mukai T., Watanabe T. K., Yamaguchi O. The genetic structure of natural populations of *Drosophila melanogaster*. XIII. Linkage disequilibrium in a large local population, *Genetics*, **77**, 771—793 (1974).

380. Mukai T., Yamaguchi O. The genetic structure of natural populations of *Drosophila melanogaster*. XI. Genetic variability in a local population, *Genetics*, **76**, 339—366 (1974).
381. Mukai T., Yamazaki T. Position effect of spontaneous mutant polygenes controlling viability in *Drosophila melanogaster*, *Proc. Japan Academy*, **40**, 840—845 (1964).
382. Mukai T., Yamazaki T. Test for selection on polymorphic isozyme genes using the population cage method, *Genetics*, **96**, 537—542 (1980).
383. Muller H. J. Variation due to change in the individual gene, *Amer. Natur.*, **56**, 32—50 (1922).
384. Muller H. J. Artificial transmutation of the gene, *Science*, **66**, 84—87 (1927).
385. Muller H. J. The gene as the basis of life, *Proc. Int. Congr. Plant Sci.*, **1**, 897—921 (1929).
386. Muller H. J. Further studies on the nature and causes of gene mutations, *Proc. Sixth Inter. Congr. Genet.*, **1**, 213—255 (1933).
387. Muller H. J. Redintegration of the symposium on genetics, paleontology, and evolution. In: *Genetics, Paleontology, and Evolution*, Jepsen G. L., Simpson G. G. and Mayr E., eds., pp. 421—445, Princeton University Press, 1949.
388. Muller H. J. Our load of mutations, *Amer. J. Hum. Genet.*, **2**, 111—176 (1950).
389. Muller H. J. The mutation theory reexamined, *Proc. X Inter. Congr. Genet.*, **1**, 306—317 (1958).
390. Muller H. J. The guidance of human evolution, *Perspectives in Biology and Medicine*, **3**(1), 1—43 (1959).
391. Muller H. J. Evolution and genetics. *Evoluzione e Genetica*, pp. 15—37, Rome, Accademia Nazionale dei Lincei, 1960.
392. Muller H. J. *Studies in Genetics*, Indiana University Press, 1962.
393. Muller H. J. The gene material as the initiator and the organizing basis of life, *Amer. Natur.*, **100**, 493—518 (1966).
394. Muller H. J. The gene material as the initiator and the organizing basis of life. In: *Heritage from Mendel*, Brink R. A., ed., pp. 417—447, University of Wisconsin Press, 1967.
395. Nagylaki T. The moments of stochastic integrals and the distribution of sojourn times, *Proc. Natl. Acad. Sci. USA*, **71**, 746—749 (1974).
396. Nagylaki T., Petes T. D. Intrachromosomal gene conversion and the maintenance of sequence homogeneity among repeated genes, *Genetics*, **110**, 315—337 (1982).
397. Narain P. A note on the diffusion approximation for the variance of the number of generations until fixation of a neutral mutant gene. *Genet. Res., Camb.*, **15**, 251—255 (1970).
398. Neel J. V. Rare variants, private polymorphisms, and locus heterozygosity in Amerindian populations, *Amer. J. Hum. Genet.*, **30**, 465—490 (1978).
399. Nei M. Gene duplication and nucleotide substitution in evolution, *Nature, Lond.*, **221**, 40—42 (1969).
400. Nei M. Genetic distance between populations, *Amer. Natur.*, **106**, 283—292 (1972).
401. Nei M. *Molecular Population Genetics and Evolution*, Amsterdam, North-Holland, 1975.
402. Nei M. The cost of natural selection and the extent of enzyme polymorphism, *Trend Biochem. Sci.*, **1**, N247—N248 (1976).
403. Nei M. Estimation of mutation rate from rare protein variants, *Amer. J. Hum. Genet.*, **29**, 225—232 (1977).
404. Nei M. Estimation of average heterozygosity and genetic distance from a small number of individuals, *Genetics*, **89**, 583—590 (1978).
405. Nei M. The theory of genetic distance and evolution of human races, *Jap. J. Hum. Genet.*, **23**, 341—369 (1978).
406. Nei M. Stochastic theory of population genetics and evolution. In: *Vivo*

- Volterra Symposium on Mathematical Models in Biology (Lecture Notes in Biomathematics 39), Barigozzi C., ed., pp. 17—47, Berlin, Springer-Verlag, 1980.
407. *Nei M., Chakraborty R.* Empirical relationship between the number of nucleotide substitutions and interspecific identity of amino acid sequences in some proteins, *J. Mol. Evol.*, **7**, 313—323 (1976).
 408. *Nei M., Chakraborty R., Fuerst P. A.* Infinite allele model with varying mutation rate, *Proc. Natl. Acad. Sci. USA*, **73**, 4164—4168 (1976).
 409. *Nei M., Feldman M. W.* Identity of genes by descent within and between populations under mutation and migration pressures, *Theoret. Pop. Biol.*, **3**, 460—465 (1972).
 410. *Nei M., Fuerst P. A., Chakraborty R.* Testing the neutral mutation hypothesis by distribution of single locus heterozygosity, *Nature, Lond.*, **262**, 491—493 (1976).
 411. *Nei M., Fuerst P. A., Chakraborty R.* Subunit molecular weight and genetic variability of proteins in natural populations, *Proc. Natl. Acad. Sci. USA*, **75**, 3359—3362 (1978).
 412. *Nei M., Imaizumi Y.* Genetic structure of human populations II. Differentiation of blood group gene frequencies among isolated human populations, *Heredity*, **21**, 183—190 (1966).
 413. *Nei M., Li W.-H.* Non-random association between electromorphs and inversion chromosomes in finite populations, *Genet. Res., Camb.*, **35**, 65—83 (1980).
 414. *Nei M., Maruyama T.* Lewontin-Krakauer test for neutral genes, *Genetics*, **80**, 395 (1975).
 415. *Nei M., Maruyama T., Chakraborty R.* The bottleneck effect and genetic variability in populations, *Evolution*, **29**, 1—10 (1975).
 416. *Nei M., Murata M.* Effective population size when fertility is inherited, *Genet. Res., Camb.*, **8**, 257—260 (1966).
 417. *Nei M., Roychoudhury A. K.* Probability of fixation and mean fixation time of an overdominant mutation, *Genetics*, **74**, 371—380 (1973).
 418. *Nei M., Roychoudhury A. K.* Sampling variances of heterozygosity and genetic distances, *Genetics*, **76**, 379—390 (1974).
 419. *Nei M., Tateno Y.* Interlocus variation of genetic distance and the neutral mutation theory, *Proc. Natl. Acad. Sci. USA*, **72**, 2758—2760 (1975).
 420. *Nevo E.* Genetic variation in natural populations: patterns and theory, *Theor. Pop. Biol.*, **13**, 121—177 (1978).
 421. *Nishioka Y., Leder A., Leder P.* Unusual α -globin-like gene that has clearly lost both globin intervening sequences, *Proc. Natl. Acad. Sci. USA*, **77**, 2806—2809 (1980).
 422. *Nichols B. P., Yanofsky Ch.* Nucleotide sequences of *trpA* of *Salmonella typhimurium* and *Escherichia coli*: An evolutionary comparison, *Proc. Natl. Acad. Sci. USA*, **76**, 5244—5248 (1979).
 423. *Novotný J.* Genealogy of immunoglobulin polypeptide chains: a consequence of amino acid interactions, conserved in their tertiary structure, *J. Theor. Biol.*, **41**, 171—180 (1973).
 424. *Nozawa K., Shotake T., Kawamoto Y., Tanabe Y.* Population genetics of Japanese monkeys: II. blood protein polymorphisms and population structure, *Primates*, **23**, 252—271 (1982).
 425. *Nozawa K., Shotake T., Okura Y.* Blood protein polymorphisms and population structure of the Japanese macaque, *Macaca fuscata fuscata*. In: *Isozymes IV. Genetics and Evolution*, Markert C. L., ed., pp. 225—241, New York, Academic Press, 1975.
 426. *Oakeshott J. G., Gibson J. B.* Is there selection by environmental ethanol on the alcohol dehydrogenase locus in *Drosophila melanogaster*? In: *Genetics Studies of Drosophila Populations*, Gibson J. B. and Oakeshott J. G., eds., pp. 103—120, Australian National University Press, 1981.

427. O'Donald P. «Haldane's dilemma» and the rate of natural selection, *Nature*, Lond., **221**, 815—816 (1969).
428. Ohba S. Genetics of Populations (in Japanese), University of Tokyo Press, 1977.
429. Ohno S. Evolution by Gene Duplication, Berlin, Springer-Verlag, 1970. [Имеется перевод: Оно С. Генетические механизмы прогрессивной эволюции. — М.: Мир, 1973.]
430. Ohno S. So much «junk» DNA in our genome. In: Brookhaven Symposia in Biology No. 23, Smith H. H., ed., pp. 366—370, New York, Gordon and Breach, 1972.
431. Ohno S. Ancient linkage groups and frozen accidents, *Nature*, Lond., **244**, 259—262 (1973).
432. Ohno S., Wolf U., Atkin N. B. Evolution from fish to mammals by gene duplication, *Hereditas*, **59**, 169—187 (1968).
433. Ohta T. Associative overdominance caused by linked detrimental mutations, *Genet. Res.*, Camb., **18**, 277—286 (1971).
434. Ohta T. Population size and rate of evolution, *J. Mol. Evol.*, **1**, 305—314 (1972).
435. Ohta T. Slightly deleterious mutant substitutions in evolution, *Nature*, Lond., **246**, 96—98 (1973).
436. Ohta T. Effect of linkage on behavior of mutant genes in a finite population, *Theor. Pop. Biol.*, **4**, 145—162 (1973).
437. Ohta T. Mutational pressure as the main cause of molecular evolution and polymorphism, *Nature*, Lond., **252**, 351—354 (1974).
438. Ohta T. Statistical analyses of *Drosophila* and human protein polymorphisms, *Proc. Natl. Acad. Sci. USA*, **72**, 3194—3196 (1975).
439. Ohta T. A simple model for treating the evolution of multigene families, *Nature*, Lond., **263**, 74—76 (1976).
440. Ohta T. Role of very slightly deleterious mutations in molecular evolution and polymorphism, *Theor. Pop. Biol.*, **10**, 254—275 (1976).
441. Ohta T. On the gene conversion model as a mechanism for maintenance of homogeneity in systems with multiple genomes, *Genet. Res.*, Camb., **30**, 89—91 (1977).
442. Ohta T. Extension to the neutral mutation random drift hypothesis. In: *Molecular Evolution and Polymorphism*, Kimura M., ed., pp. 148—167, Mishima, National Institute of Genetics, 1977.
443. Ohta T. Theoretical study on genetic variation in multigene families, *Genet. Res.*, Camb., **31**, 13—28 (1978).
444. Ohta T. Theoretical population genetics of repeated genes forming a multigene family, *Genetics*, **88**, 845—861 (1978).
445. Ohta T. An extension of a model for the evolution of multigene families by unequal crossing-over, *Genetics*, **91**, 591—607 (1979).
446. Ohta T. Linkage disequilibrium between amino acid sites in immunoglobulin genes and other multigene families, *Genet. Res.*, Camb., **36**, 181—197 (1980).
447. Ohta T. Evolution and Variation of Multigene Families (Lecture Notes in Biomathematics 37), Berlin, Springer-Verlag, 1980.
448. Ohta T. Genetic variation in small multigene families, *Genet. Res.*, Camb., **37**, 133—149 (1981).
449. Ohta T. Population genetics of selfish DNA, *Nature*, Lond., **292**, 648—649 (1981).
450. Ohta T. Linkage disequilibrium due to random drift in finite subdivided populations, *Proc. Natl. Acad. Sci. USA*, **79**, 1940—1944 (1982).
451. Ohta T. Allelic and non-allelic homology of a supergene family, *Proc. Natl. Acad. Sci. USA*, **79** (1982), in press.
452. Ohta T., Kimura M. Development of associative overdominance through linkage disequilibrium in finite populations, *Genet. Res.*, Camb., **16**, 165—177 (1970).

453. Ohta T., Kimura M. Functional organization of genetic material as a product of molecular evolution, *Nature, Lond.*, **233**, 118—119 (1971).
454. Ohta T., Kimura M. On the constancy of the evolutionary rate of cistrons, *J. Mol. Evol.*, **1**, 18—25 (1971).
455. Ohta T., Kimura M. Linkage disequilibrium between two segregating nucleotide sites under the steady flux of mutations in a finite population, *Genetics*, **68**, 571—580 (1971).
456. Ohta T., Kimura M. Behavior of neutral mutants influenced by associated overdominant loci in finite populations, *Genetics*, **69**, 247—260 (1971).
457. Ohta T., Kimura M. A model of mutation appropriate to estimate the number of electrophoretically detectable alleles in finite a population, *Genet. Res., Camb.*, **22**, 201—204 (1973).
458. Ohta T., Kimura M. Some calculations on the amount of selfish DNA, *Proc. Natl. Acad. Sci. USA*, **78**, 1129—1132 (1981).
459. Omoto K. The genetic origins of Philippine Negritos, *Current Anthropology*, **22**, 421—422 (1981).
460. Omoto K. Allelic diversity of human populations in the Asian-Pacific area. In: *Molecular Evolution, Protein Polymorphism and the Neutral Theory*, Kimura M., ed., pp. 193—214. Tokyo, Japan Scientific Societies Press/Berlin, Springer-Verlag, 1982.
461. Orgel L. E., Crick F. H. C. Selfish DNA: the ultimate parasite, *Nature, Lond.*, **284**, 604—607 (1980).
462. Orgel L. E., Crick F. H. C., Spienza C. Selfish DNA, *Nature, Lond.*, **288**, 645—646 (1980).
463. Ozeki H., Inokuchi H., Yamao F., Kodaira M., Sakano H., Ikemura T., Shimura Y. Genetics of nonsense suppressor tRNAs in *Escherichia coli*. In: *Transfer RNA: Biological Aspects*, Söll D., Abelson J. and Schimmel P., eds., pp. 341—362, Cold Spring Harbor Laboratory, 1980.
464. Pearson K. Walter Frank Raphael Weldon, 1860—1906, *Biometrika*, **5**, 1—52 (1906).
465. Pearson K. *Tables of the Incomplete Beta-Function*, Cambridge University Press, 1934.
466. Perier F., Efstratiadis A., Lomedico P., Gilbert W., Kolodner R., Dodgson J. The evolution of genes: the chicken preproinsulin gene, *Cell*, **20**, 555—566 (1980).
467. Perutz M. F., Lehmann H. Molecular pathology of human haemoglobin, *Nature, Lond.*, **219**, 902—909 (1968).
468. Pilbeam D. New hominoid skull material from the Miocene of Pakistan, *Nature*, **295**, 232—234 (1982).
469. Post L. E., Strycharz G. D., Nomura M., Lewis H., Dennis P. P. Nucleotide sequence of the ribosomal protein gene cluster adjacent to the gene for RNA polymerase subunit β in *Escherichia coli*, *Proc. Natl. Acad. Sci. USA*, **76**, 1697—1701 (1979).
470. Powell J. R. Genetic polymorphisms in varied environments, *Science*, **174**, 1035—1036 (1979).
471. Powell J. R. Isozymes and non-Darwinian evolution: a re-evaluation. In: *Isozymes IV Genetics and Evolution*, Markert C. L., ed., pp. 9—26, New York, Academic Press, 1975.
472. Prakash S., Lewontin R. C. A molecular approach to the study of genic heterozygosity in natural populations, III. Direct evidence of coadaptation in gene arrangements of *Drosophila*, *Proc. Natl. Acad. Sci. USA*, **59**, 398—405 (1968).
473. Prakash S., Lewontin R. C., Hubby J. L. A molecular approach to the study of genic heterozygosity in natural populations. IV. Patterns of genic variation in central, marginal and isolated populations of *Drosophila pseudoobscura*, *Genetics*, **61**, 841—858 (1969).
474. Proudfoot N. Pseudogenes, *Nature, Lond.*, **286**, 840—841 (1980).
475. Proudfoot N. J., Maniatis T. The structure of a human α -globin pseudogene and its relationship to α -globin gene duplication, *Cell*, **21**, 537—544 (1980).

476. *Prout T.* The estimation of fitnesses from genotypic frequencies, *Evolution*, **19**, 546—551 (1965).
477. *Provine W. B.* The Origins of Theoretical Population Genetics, The University of Chicago Press, 1971.
478. *Racine R. R., Langley C. H.* Genetic heterozygosity in a natural population of *Mus musculus* assessed using two-dimensional electrophoresis, *Nature*, Lond., **283**, 855—857 (1980).
479. *Ramshaw J. A. M., Coyne J. A., Lewontin R. C.* The sensitivity of gel electrophoresis as a detector of genetic variation, *Genetics*, **93**, 1019—1037 (1979).
480. *Reeder R. H., Brown D. D., Wellauer P. K., Dawid I. B.* Patterns of ribosomal DNA spacer lengths are inherited, *J. Mol. Biol.*, **105**, 507—516 (1976).
481. *Rensch B.* Evolution above the Species Level, Columbia University Press, 1960.
482. *Richards R. I., Shine J., Ullrich A., Wells J. R. I., Goodman H. M.* Molecular cloning and sequence analysis of adult chicken β globin cDNA, *Nucl. Acids Res.*, **7**, 1137—1146 (1979).
483. *Richardson J. S.* The anatomy and taxonomy of protein structure, *Advances in Protein Chemistry*, **34**, 168—339 (1981).
484. *Richardson R. H., Richardson M. E., Smouse P. E.* Evolution of electrophoretic mobility in the *Drosophila mulleri* complex. In: *Isozyme*, vol. 4, Genetics and Evolution, Markdert C. L., ed., pp. 533—545, New York, Academic Press, 1975.
485. *Robertson A.* The effect of selection against extreme deviants based on deviation or on homozygosis, *J. Genet.*, **54**, 236—248 (1956).
486. *Robertson A.* Inbreeding in artificial selection programmes, *Genet. Res., Camb.*, **2**, 189—194 (1961).
487. *Robertson A.* Selection for heterozygotes in small populations, *Genetics*, **47**, 1291—1300 (1962).
488. *Robertson A.* The nature of quantitative genetic variation. In: *Heritage from Mendel*, Brink R. A., ed., pp. 265—280, University of Wisconsin Press, 1967.
489. *Robertson A.* The spectrum of genetic variation. In: *Population Biology and Evolution*, Lewontin R. C., ed., pp. 5—16, Syracuse University Press, 1968.
490. *Robertson A.* The reduction in fitness from genetic drift at heterotic loci in small populations, *Genet. Res., Camb.*, **15**, 257—259 (1970).
491. *Robertson A.* Remarks on the Lewontin-Krakauer test, *Genetics*, **80**, 396 (1975).
492. *Romer A. S.* Major steps in vertebrate evolution, *Science*, **158**, 1629—1637 (1967).
493. *Romer A. S.* The Procession of Life, London, Weidenfeld and Nicholson, 1968.
494. *Romero-Herrera A. E., Lieska N., Nasser S.* Characterization of the myoglobin the lamprey *Petromyzon marinus*, *J. Mol. Evol.*, **14**, 259—266 (1979).
495. *Roughgarden J.* Theory of Population Genetics and Evolutionary Ecology: An Introduction, New York, Macmillan, 1979.
496. *Ruffié J.* De la Biologie a la Culture, Paris, Flammarion, 1976.
497. *Ruffié J.* Traité du Vivant, Paris, Fayard, 1982.
498. *Sakaizumi M.* Biochemical genetics of *Oryzias* populations. Doctoral Dissertation, Faculty of Science, University of Tokyo, 1982.
499. *Sakaizumi M., Egami N., Moriwaki K.* Allozymic variation in wild populations of the fish, *Oryzias latipes*, *Proc. Japan Acad. Ser. B*, **56**, 448—451 (1980).
500. *Sakano H., Rogers J. H., Hüppi K., Brack C., Traunecker A., Maki R., Wall R., Tonegawa S.* Domains and the hinge regions of an immunoglobulin

- heavy chain are encoded in separate DNA segments, *Nature*, Lond., **277**, 627—633 (1979).
501. *Salser W.* Globin mRNA sequences: analysis of base pairing and evolutionary implication, Cold Spring Harbor Symp. Quant. Biology, **42** («Chromatin», part 2), 985—1002 (1978).
 502. *Sarich V. M., Wilson A. C.* Immunological time scale for hominid evolution, *Science*, **158**, 1200—1202 (1967).
 503. *Schaffner W., Kunz G., Daetwyler H., Telford J., Smith H. O., Birnstiel M. L.* Genes and spacers of cloned sea urchin histone DNA analyzed by sequencing, *Cell*, **14**, 655—671 (1978).
 504. *Schopf T. J. M.* Patterns of evolution: a summary and discussion. In: *Patterns of Evolution*, Hallam A., ed., pp. 547—561, Amsterdam, Elsevier, 1977.
 505. *Seeburg P. H., Shine J., Martial J. A., Baxter J. D., Goodman H. M.* Nucleotide sequence and amplification in bacteria of structural gene for rat growth hormone, *Nature*, Lond., **270**, 486—494 (1977).
 506. *Selander R. K.* Genic variation in natural populations. In: *Molecular Evolution*, Ayala F. J., ed., pp. 21—45, Sunderland, Sinauer, 1976.
 507. *Selander R. K., Levin B. R.* Genetic diversity and structure in *Escherichia coli* populations, *Science*, **210**, 545—547 (1980).
 508. *Shaw C. S.* How many genes evolve? *Biochemical Genetics*, **4**, 275—283 (1970).
 509. *Shibata S., Iuchi I., Miyaji T.* Abnormal hemoglobins discovered in Japan, *Acta Haematologica Japonica*, **29**, 115—127 (1966).
 510. *Simmons M. J., Crow J. F.* Mutations affecting fitness in *Drosophila* populations, *Annu. Rev. Genet.*, **11**, 49—78 (1977).
 511. *Simons E. L.* Man's immediate forerunners, *Philosophical Transactions of the Royal Society (London)*, vol. B, **292**, p. 21—41 (1981).
 512. *Simpson G. G.* Tempo and Mode in Evolution, Columbia University Press, 1944. [Имеется перевод: Симпсон Дж. Темпы и формы эволюции. — М.: ИЛ, 1948.]
 513. *Simpson G. G.* The Meaning of Evolution, Yale University Press, 1949.
 514. *Simpson G. G.* The Major Features of Evolution, Columbia University Press, 1953.
 515. *Simpson G. G.* This View of Life, New York, Harcourt, Brace and World, 1964.
 516. *Singh R. S., Lewontin R. C., Felton A. A.* Genic heterogeneity within electrophoretic «alleles» of xanthine dehydrogenase in *Drosophila pseudoobscura*, *Genetics*, **84**, 609—629 (1976).
 517. *Slatkin M.* On treating the chromosome as the unit of selection, *Genetics*, **72**, 157—168 (1972).
 518. *Slatkin M.* Gene flow and genetic drift in a species subject to frequent local extinctions, *Theor. Pop. Biol.*, **12**, 253—262 (1977).
 519. *Smith G. P.* Unequal crossover and the evolution of multigene families, Cold Spring Harbor Symp. Quant. Biol., **38**, 507—513 (1974).
 520. *Sneath P. H. A.* Relations between chemical structure and biological activity in peptides, *J. Theor. Biol.*, **12**, 157—195 (1966).
 521. *Sneath P. H. A.* The estimation of differences in protein evolution rates, *Proc. Geol. Assoc.*, **91** (1 and 2), 71—79 (1980).
 522. *Somero G. N., Soulé M.* Genetic variation in marine fishes as a test of the niche-variation hypothesis, *Nature*, Lond., **249**, 670—672 (1974).
 523. *Soulé M.* Allozyme variation: its determinants in space and time. In: *Molecular Evolution*, Ayala F. J., ed., pp. 60—77, Sunderland, Sinauer, 1976.
 524. *Spiess E. B.* Genes in Populations, New York, John Wiley and Sons, 1977.
 525. *Spiess E. B., Allen A. C.* Release of genetic variability through recombination. VII. Second and third chromosomes of *Drosophila melanogaster*, *Genetics*, **46**, 1531—1553 (1961).

526. Spieth P. T. Population genetics of allozyme variation in *Neurospora intermedia*, Genetics, **80**, 785—805 (1975).
527. Spofford J. B. Heterosis and the evolution of duplications, Amer. Natur., **103**, 407—432 (1969).
528. Stamatoyannopoulos G., Chen S.-H., Fukui M. Liver alcohol dehydrogenase in Japanese: high population frequency of atypical form and its possible role in alcohol sensitivity, Amer. J. Hum. Genet., **27**, 789—796 (1975).
529. Stebbins G. L., Lewontin R. C. Comparative evolution at the levels of molecules, organisms and populations. In: Proc. of the Sixth Berkeley Symp. on Math. Stat. and Prob. vol. V: Darwinian, Neo-Darwinian, and Non-Darwinian Evolution, LeCam L. M., Neyman J. and Scott E. L., eds., pp. 23—42, University of California Press, 1972.
530. Steele E. J. Somatic Selection and Adaptive Evolution, Toronto, Williams-Wallace, 1979.
531. Stenzel P. Opossum Hb chain sequence and neutral mutation theory, Nature, Lond., **252**, 62—63 (1974).
532. Stewart F. M. Variability in the amount of heterozygosity maintained by neutral mutations, Theor. Pop. Biol., **9**, 188—201 (1976).
533. Sturtevant A. H. A History of Genetics, New York, Harper and Row, 1965.
534. Subak-Sharpe J. H., Bürk R. R., Crawford L. V., Morrison J. M., Hay J., Keir H. M. An approach to evolutionary relationships of mammalian DNA viruses through analysis of the pattern of nearest neighbor base sequences, Cold Spring Harbor Sump. Quant. Biol., **31**, 737—748 (1966).
535. Sures I., Goettel D. V., Gray A., Ullrich A. Nucleotide sequence of human preproinsulin complementary DNA, Science, **208**, 57—59 (1980).
536. Sures I., Lowry J., Kedes L. H. The DNA sequence of sea urchin (*S. purpuratus*) H2A, H2B and H3 histone coding and spacer regions, Cell, **15**, 1033—1044 (1978).
537. Sved J. A. Possible rates of gene substitution in evolution, Amer. Natur., **102**, 283—293 (1968).
538. Sved J. A. The relationship between genotype and fitness for heterotic models. In: Population Genetics and Ecology, Karlin S. and Nevo E., eds., pp. 441—463; New York, Academic Press, 1976.
539. Sved J. A., Reed T. E., Bodmer W. F. The number of balanced polymorphisms that can be maintained in a natural population, Genetics, **55**, 469—481 (1967).
540. Takahata N. Composite stepwise mutation model under the neutral mutation hypothesis, J. Mol. Biol., **15**, 13—20 (1980).
541. Takahata N., Kimura M. Genetic variability maintained in a finite population under mutation and autocorrelated random fluctuation of selection intensity, Proc. Natl. Acad. Sci. USA, **76**, 5813—5817 (1979).
542. Takahata N., Kimura M. A model of evolutionary base substitutions and its application with special reference to rapid change of pseudogenes, Genetics, **98**, 641—657 (1981).
543. Takahata N., Maruyama T. Polymorphism and loss of duplicate gene expression: a theoretical study with application to tetraploid fish, Proc. Natl. Acad. Sci. USA, **76**, 4521—4525 (1979).
544. Tartof K. Unequal mitotic sister chromatid exchange and disproportionate replication as mechanisms regulating ribosomal RNA gene redundancy, Proc. Cold Spring Harbor Symp. Quant. Biol., **37**, 491—500 (1974).
545. Tateno Y., Nei M., Goodman M. et al's method for augmenting the number of nucleotide substitutions, J. Mol. Evol., **11**, 67—73 (1978).
546. Temin R. G., Meyer H. U., Dawson P. S., Crow J. F. The influence of epistasis of homozygous viability depression in *Drosophila melanogaster*, Genetics, **61**, 497—519 (1969).
547. Thoday J. M. Non-Darwinian «evolution» and biological progress, Nature, Lond., **255**, 675—677 (1975).

548. *Thoday J. M.* Wrong assumption, *Trends in Biochemical Sciences*, **1**, N247 (1976).
550. *Thompson J. N., Kaiser T. N.* Selection acting upon slow-migrating ADH alleles differing in enzyme activity, *Heredity*, **38**, 191—195 (1977).
551. *Tomita S., Riggs A.* Studies of the interaction of 2,3-diphosphoglycerate and carbon dioxide with hemoglobins from mouse, man and elephant, *J. Biol. Chem.*, **246**, 547—554 (1971).
552. *Tsujimoto Y., Hirose S., Tsuda M., Suzuki Y.* Promotor sequence of fibroin gene assigned by in vitro transcription system, *Proc. Natl. Acad. Sci. USA*, **78**, 4838—4842 (1981).
553. *Valentine J. W.* Genetic strategies of adaptation. In: *Molecular Evolution*, Ayala F. J., ed., pp. 78—94, Sunderland, Sinauer, 1976.
554. *Valenzuela P., Quiroga M., Zaldivar J., Rutter W. J., Kirschner M. W., Cleveland D. W.* Nucleotide and corresponding amino acid sequences encoded by α and β tubulin mRNAs, *Nature, Lond.*, **289**, 650—655 (1981).
555. *van Delden W., Boerema A. C., Kamping A.* The alcohol dehydrogenase polymorphism in populations of *Drosophila melanogaster*. I. Selection in different environments, *Genetics*, **90**, 161—191 (1978).
556. *van der Berg J., van Ooyen A., Mantei N., Schamböck A., Grosveld G., Flavell R. A., Weissmann Ch.* Comparison of closed rabbit and mouse β -globin genes showing strong evolutionary divergence of two homologous pairs of introns, *Nature, Lond.*, **276**, 37—44 (1978).
557. *Vanin E. F., Goldberg G. I., Tucker P. W., Smithies O.* A mouse α -globin-related pseudogene lacking intervening sequences, *Nature, Lond.*, **286**, 222—226 (1980).
558. *van Ooyen A., van den Berg J., Mantei N., Weissman C.* Comparison of total sequence of a cloned rabbit β -globin gene and its flanking regions with a homologous mouse sequence, *Science*, **206**, 337—344 (1979).
559. *van Roode J. H. G., Orgel L. E.* Template-directed synthesis of oligoguanylates in the presence of metal ions, *J. Mol. Biol.*, **144**, 579—585 (1980).
560. *Vogel F.* Point mutations and human hemoglobin variants, *Humangenetik*, **8**, 1—26 (1969).
561. *Vogel F., Motulsky A. G.* *Human Genetics*, Berlin, Springer-Verlag, 1979.
562. *Waddington C. H.* *The Strategy of the Genes*, London, George Allen and Unwin, 1957.
563. *Wallace B.* The average effect of radiation-induced mutations on viability in *Drosophila melanogaster*, *Evolution*, **12**, 532—552 (1958).
564. *Wallace B.* *Topics in Population Genetics*, New York, Norton, 1968.
565. *Ward R. D.* Relationship between enzyme heterozygosity and quaternary structure, *Biochem. Genet.*, **15**, 123—135 (1977).
566. *Ward R. D., Beadmore J. A.* Protein variation in the plaice, *Pleuronectes platessa* L., *Genet. Res., Camb.*, **30**, 45—62 (1977).
567. *Washburn, S.* The evolution of man, *Scientific American*, **239** (3), 146—154 (1978).
568. *Wasylyk B., Derbyshire R., Guy A., Molko D., Roget A., Téoule R., Chambon P.* Specific in vitro transcription of conalbumin gene is drastically decreased by single-point mutation in T-A-T-A box homology sequence, *Proc. Natl. Acad. Sci. USA*, **77**, 7024—7028 (1980).
569. *Watson J. D.* *Molecular Biology of the Gene*, New York, Benjamin, 1970. [Имеется перевод: Уотсон Дж. Молекулярная биология гена. — М.: Мир, 1978.]
570. *Watterson G. A.* Some theoretical aspects of diffusion theory in population genetics, *Annu. Math. Stat.*, **33**, 939—957 (1962).
571. *Watterson G. A.* On the number of segregating sites in genetical models without recombination, *Theor. Pop. Biol.*, **7**, 256—276 (1975).
572. *Watterson G. A.* Heterosis or neutrality? *Genetics*, **85**, 789—814 (1977).

573. Watterson G. A. Reversibility and the age of an allele II. Two-allele models, with selection and mutation, Theor. Pop. Biol., **12**, 179—196 (1977).
574. Watterson G. A. An analysis of multi-allele data, Genetics, **88**, 171—179 (1978).
575. Watterson G. A. The homozygosity test of neutrality, Genetics, **88**, 405—417 (1978).
576. Weinberg W. Über den Nachweis der Vererbung beim Menschen, Jahresh. Verein f. Vaterl. Naturk. Württem., **64**, 368—382 (1908).
577. Weir B. S., Brown A. H. D., Marshall D. R. Testing for selective neutrality of electrophoretically detectable protein polymorphisms, Genetics, **84**, 639—659 (1976).
578. Weldon W. F. R. A first study of natural selection in *Clausilia laminata* (Montagu), Biometrika, **1**, 109—124 (1901).
579. White T. J., Mross G. A., Osserman E. F., Wilson A. C. Primary structure of rat lysozyme, Biochemistry, **16**, 1430—1436 (1977).
580. Wills C. The mutational load in two natural populations of *Drosophila pseudoobscura*, Genetics, **53**, 281—294 (1966).
581. Wills C. Rank-order selection is capable of maintaining all genetic polymorphisms, Genetics, **89**, 403—417 (1978).
582. Wills C. Genetic Variability, Oxford University Press, 1981.
583. Wilson E. O. Sociobiology. The Belknap Press of Harvard University Press, 1975.
584. Wilson A. C., Carlson S. S., White T. J. Biochemical evolution, Annu. Rev. Biochem., **46**, 573—639 (1977).
585. Wright S. Systems of mating, Genetics, **6**, 111—178 (1921).
586. Wright S. Evolution in Mendelian populations, Genetics, **16**, 97—159 (1931).
587. Wright S. The roles of mutation, inbreeding, crossbreeding and selection in evolution, Proc. VI. Intern. Congr. Genet., **1**, 356—366 (1932).
588. Wright S. Physiological and evolutionary theories of dominance, Amer. Natur., **68**, 25—53 (1934).
589. Wright S. The analysis of variance and the correlations between relatives with respect to deviations from an optimum, J. Genet., **30**, 243—256 (1935).
590. Wright S. The distribution of gene frequencies in populations. Proc. Natl. Acad. Sci. USA, **23**, 307—320 (1937).
591. Wright S. The distribution of gene frequencies under irreversible mutation, Proc. Natl. Acad. Sci. USA, **24**, 253—259 (1938).
592. Wright S. Size of population and breeding structure in relation to evolution, Science, **87**, 430—431 (1938).
593. Wright S. Statistical genetics and evolution, Bull. Amer. Math. Soc., **48**, 223—246 (1942).
594. Wright S. Isolation by distance, Genetics, **28**, 114—138 (1943).
595. Wright S. The differential equation of the distribution of gene frequencies, Proc. Natl. Acad. Sci. USA, **31**, 382—389 (1945).
596. Wright S. Isolation by distance under diverse systems of mating, Genetics, **31**, 39—59 (1946).
597. Wright S. Adaptation and selection. In: Genetics, Paleontology, and Evolution, Jepsen G. L., et al., eds., pp. 365—389, Princeton University Press, 1949.
598. Wright S. Fisher and Ford on «the Sewall Wright effect», American Scientist, **39**, 452—458 (1951).
599. Wright S. The genetical structure of populations, Ann. Eugen., **15**, 323—354 (1951).
600. Wright S. Polyallelic random drift in relation to evolution, Proc. Natl. Acad. Sci. USA, **55**, 1074—1081 (1966).
601. Wright S. Evolution and the Genetics of Populations, vol. 2, The Theory of Gene Frequencies, University of Chicago, 1969.

602. Wright S. Random drift and the shifting balance theory of evolution. In: Mathematical Topics in Population Genetics, Kojima K., ed., pp. 1—31, Berlin, Springer-Verlag, 1970.
603. Wright S. Evolution and the Genetics of Populations, vol. 3, Experimental Results and Evolutionary Deductions, University of Chicago Press, 1977.
604. Wright S. Evolution and the Genetics of Populations, vol. 4, Variability Within and Among Natural Populations, University of Chicago Press, 1978.
605. Wright S., Dobzhansky Th. Genetics of natural populations. XII. Experimental reproduction of some of the changes caused by natural selection in certain populations of *Drosophila pseudoobscura*, Genetics, 31, 125—150 (1946).
606. Yamazaki T. Measurement of fitness at the esterase-5 locus in *Drosophila pseudoobscura*, Genetics, 67, 579—603 (1971).
607. Yamazaki T. Enzyme polymorphism and functional difference: mean, variance, and distribution of heterozygosity. In: Molecular Evolution and Polymorphism, Kimura M., ed., pp. 127—147, Mishima, National Institute of Genetics, 1977.
608. Yamazaki T. Genetics variabilities in natural population of haploid plant, *Conocephalum conicum*. I. The amount of heterozygosity, Jap. J. Genet., 56, 373—383 (1981).
609. Yamazaki T., Maruyama T. Evidence for the neutral hypothesis of protein polymorphism, Science, 178, 56—58 (1972).
610. Yamazaki T., Maruyama T. Evidence that enzyme polymorphism are selectively neutral, Nature, New Biol., 245, 140—141 (1973).
611. Yamazaki T., Maruyama T. Evidence that enzyme polymorphisms are selectively neutral, but blood group polymorphisms are not, Science, 183, 1091—1092 (1974).
612. Yamazaki T., Maruyama T. Isozyme polymorphism maintenance mechanisms viewed from the standpoint of population genetics. In: *Isozymes IV*. Genetics and Evolution, Markert C. L., ed., pp. 103—114, New York, Academic Press, 1975.
613. Yanase T. Basis of medical genetics — Molecular disease Part 6. Clinic All-Round (*Sōgō Rinshō*), 20 (12), 2964—2976 (in Japanese) (1971).
614. Yanofsky C. Gene structure and protein structure, Scientific American, 216 (5), 80—82 (1967).
615. Yanofsky C., Horn V., Thorpe D. Protein structure relationships revealed by mutational analysis, Science, 146, 1593—1594 (1964).
616. Yarbrough K., Kojima K. The mode of selection at the polymorphic esterase 6 locus in cage populations of *Drosophila melanogaster*, Genetics, 57, 677—686 (1967).
617. Yoshimaru H., Mukai T. Lack of experimental evidence for frequency-dependent selection at the alcohol dehydrogenase locus in *Drosophila melanogaster*, Proc. Natl. Acad. Sci. USA, 76, 876—878 (1979).
618. Zimmer E. A., Martin S. L., Beverley S. M., Kan Y. W., Wilson A. C. Rapid duplication and loss of genes coding for the α chains of hemoglobin, Proc. Natl. Acad. Sci. USA, 77, 2158—2162 (1980).
619. Zouros E. Hybrid molecules and the superiority of the heterozygote, Nature, Lond., 262, 227—229 (1976).
620. Zuckerkandl E. Evolutionary process and evolutionary noise at the molecular level II. A selectionist model for random fixations in proteins, J. Mol. Evol., 7, 269—311 (1976).
621. Zuckerkandl E. Multilocus enzymes, gene regulation, and genetic sufficiency, J. Mol. Evol., 12, 57—89 (1978).
622. Zuckerkandl E., Pauling L. Evolutionary divergence and convergence in proteins. In: Evolving Genes and Proteins, Bryson V. and Vogel H. J., eds., pp. 97—166, New York, Academic Press, 1965.

Предметный указатель

- Австралопитек** 79
Адаптивный полиморфизм 33
Аденин 39, 217
Акула, гемоглобин 88, 90
Аланин 40
Алкогольдегидрогеназа 292—295, 337, 353
Аллель (и) 230, 231
— гетерозисный 45
— множественный 33
— мутантный 171—175
— нейтральный 333—338, 361
— определяющий пол 243
— редкий 260—262, 329—333
— сверхдоминантный 44
— селективно эквивалентные 51—53
— часто встречающийся 260—262
Аллозимный локус 288, 292, 301
Аминокислоты гидрофильные 187
— гидрофобные 186
Амфибии 75
Аргинин 40, 190
Аскорбиновая кислота 139
Аспарагин 40
Аспарагиновая кислота 40
Африканские человекообразные 80
- Балмера правило** 262
Барсук, гемоглобин 99
Бесчелюстные 74—75, 81, 108
Биометрия 19
Бора эффект 126
Брахидактилия у человека 20
Бурый лемур, гемоглобин 92
Бык, гемоглобин 99
— миоглобин 99
— рибонуклеаза 99
- Валин** 40
Вегенера гипотеза дрейфа континентов 76
Вероятность фиксации гена 39, 126, 172
— единичной мутации 62, 63, 66
Виртуальная гетерозиготность 299
- Внутрихромосомный неравный кроссинговер** 135, 352
«Время пребывания» 255
Вставные последовательности 49
- Гаплоидные организмы** 298—303
Гармоническое среднее 58, 60
Гегенбауэра полнном 245
Гемовый «карман» 186, 187, 348
Гемоглобин 87—108, 116, 136, 185—188, 197—198, 208, 348
— варианты 187
Ген, интронно-экзонная структура 359
— репрессор *lac*-оперона 214
Генетическая изменчивость 320—330
— — на молекулярном уровне 284—338
«Генетический хлам» 69
Генетическое расстояние 280—283
Генный полиморфизм 299
Гетерозигота 32—35
— по двум летелям 35
— по одной летели 35
— преимущество 298, 353
— свободная от летелей 35
Гетерозиготность 60—71, 306—359
— виртуальная 60, 298—300
— дисперсия 306, 307, 323—325, 353
— распределение 285—289
— средняя 323, 324, 326
Гетерозис при инверсионном полиморфизме 32
Гетерозисные аллели 45
Гетеромультимер 321
Гиббон 136
— гемоглобин 92
Гиллеси «модель со случайной средой» 71
Гипотеза изменчивости ниш 318
— мутационно-дрейфовая, или нейтральности 41—50
— нейтральных мутаций 123, 304
— «полной нейтральности» 304, 306

— постоянства скорости эволюции 123, 130, 347
 — сверхдоминирования 314—315
 — стабильности трофических ресурсов 319, 355
 Гистоны 94, 184—205, 342, 350
 Гистидин 40, 187, 208
 Глиции 40, 190
 α -Глобин, мутация 125
 Глутамин 40
 Глутаминовая кислота 40
 Гомеостаз 69
 — генетический 32, 38
 Гомозигота 33, 35, 44, 49
 — по инверсии 32
 Горилла 136
 — гемоглобин 92
 Гремучая змея, цитохром *c* 99, 122
Грэнсема геномная гипотеза 216
 Груз замещений, или субституционный 41—44, 162
 — генетический 154—163
 — дисметрический 160—161
 — обусловленный дрейфом 160—161
 — сегрегационный 45, 156, 158, 326
 Гуанин 40, 217

Давление мутаций 51
 Дарвин как единица скорости эволюции 85, 95
Дарвина теория естественного отбора 16—17
 Дарвиновский отбор 114, 183, 184, 191, 204
 Двухфакторная седловина 28
 Дем 229, 288, 333
 Детерминистическая теория 26, 257
Дикхьюзена — *Хартла* эффект 303, 361
 Динозавр 76, 78
Дирака дельта-функция 255
 Дисперсия 54, 58, 59, 98, 244, 283
 — гетерозиготности 306, 307, 324—326, 353
 ДНК 185, 204—207, 340, 341, 350
 — метилирование 72
 — нуклеотидный состав 72
 — репликация 278
Добржанского «балансовая» гипотеза 32—34, 152
 — «классическая» гипотеза 32, 33
 — модель 141—142
 Доменный перенос 360
 Дрозофила, инверсионный полиморфизм 32—33
 — мутации 24
 — хромосома *CIB* 31

Дупликация гена 129
 — датировка 108
Дэйхоф, матрица вероятностей мутирования 96

Европеоидная раса 283
 Единица молекулярной эволюции 95
 Естественный отбор 70
 — значение 143—145
 — определение 143
 — отрицательный 70, 73, 144
 — положительный 144
 Ехидна, гемоглобин 77, 78

Замена мутационная 64
 — аминокислот 89, 90, 96—98, 104, 105, 121, 343
 — — скорость накопления в ходе эволюции 39, 90, 93—95, 97
 — гена 62—63
 — молчащая (синонимичная) 49, 118, 119, 197—207, 342, 347, 350
 — мутационная 53, 102, 113, 171—183
 — нуклеотидов, скорость 39, 120, 211
 — — по типу транзиции 113, 114, 120, 198
 — — — трансверсии 113, 114, 120, 198

Изменение частоты гена 251—259
 Изоаллели 45
 Изолейцин 40
 Изоферментный локус 293, 296, 297, 317, 353
 Иммуноглобулины 136
 Инсулин 95, 138—140
 Интегрированный коадаптированный комплекс 38
 Интроны 49, 120, 129, 201, 208, 342
Йогансена теория чистых линий 19

Карликовый шимпанзе 136
 Карп, гемоглобин 88, 91
 Кенгуру 88
Кимуры диффузионные модели 23
 — и *Кроу* метод 165, 166
 — — формула 203—204, 306
 Кистеперые рыбы 75
 Кит, гемоглобин 99
 — рибонуклеаза 99
Кларка и *Аллендорфа* модель 298
 Коадаптированный генный комплекс 32

- Коата, гемоглобин 92
Кодоны 40, 65, 72, 107—125, 187, 200, 210, 212, 226, 281, 282, 349, 350
— использование 212
— синонимичные 211—221, 363
— третье положение 50, 115, 176, 182, 202
Колмогорова уравнение 133
Конверсия гениая 209
Копытные 84
Коэффициент гибридинга 334—335
— отбора 62, 66, 71, 123—157, 345, 360
— различия 177, 178
Кривая «доза — эффект» 24
Кролик, гемоглобин 97, 99, 116, 197
Кроссинговер 135, 352
Крыса, рибонуклеаза 99
Ксайтиндегидрогеназный локус 290
Курица, гемоглобин 88, 90, 91, 116
- α-Лактальбумин** 129, 130
Левонтина и *Кракауэра* тест 309—310
Лейции 40, 212—216
Лизин 40
Лизоцим 95, 129
Логарифм иммунологического расстояния 80
Лошадь, гемоглобин 97, 99
— миоглобин 99
— рибонуклеаза 99
- Максама — Гилберта** метод 110, 112
«Мембранин» 328
Мендель, законы 18
Метаотбор 36
Мечехвост 286, 357
Метионин 40
Метод диффузионных уравнений 42, 67, 136
— максимального правдоподобия 281
— наибольшей экономии 109, 111
— стохастический 22, 340
— филетического расстояния 97
— электрофоретический 39, 44, 340
Миоглобин 87, 93, 99, 103, 107, 108
Михаэлиса константа 302
Млекопитающие, миоглобин 180
— появление 77—78
Множественные аллели 33
Молекул структура 184—198
— — влияние на генетическую изменчивость 320—329
Молекулярная эволюция 121—142
Модель аллельных мутаций 230—243
— бесконечного числа аллелей 230—264, 306, 325, 356
— — сайтов 64, 65, 267—271
— варьирующей скорости мутирования 307
— «гиперцикла» 360
— диффузионная 23, 38, 247
— со случайной средой 71
— «степпинг-стоун» 249, 333
— ступенчатого мутирования 260—267
— тороидального пространства 335—336
— трех типов замен 198
— эффективно нейтральных мутаций 271—280
Молекулярная эволюция, принципы 121, 127, 138
— — стандартная скорость 95
Молекулярные эволюционные часы 105, 121
Монголоидная раса 283
Мономерный фермент 320, 322, 323, 355
Монотормизм 226, 286, 290, 297, 331, 358
Монте-Карло метод 55, 110, 251, 310, 325
Морская свинья, инсулин 138—140
Мультимерный фермент 320, 322, 323, 355
Мутантный аллель вредный 173
— — гетерозисный 45
— — селективно благоприятный 68
— — селективно нейтральный 51, 67, 102, 121
— — с отрицательным сверхдоминированием 173
Мутации множественные 108—109
Мутация(и) благоприятная 43, 66
— вредная 43
— гениая 53
— нейтральная 61—64, 65—66
— почти нейтральная 61—64
— селективно эквивалентные 69
— электрофоретически различимые 332
Мэйнарда Смита пороговая модель 47—48
Мышь, гемоглобин 97, 99, 126
- Наследование приобретенных признаков 16
Насекомые 76
Недарвиновская эволюция 45, 46
Негроидная раса 283
Неоклассическая теория 46
Неоламаркизм 15

Нея тест 306—310, 253

Нейтральный мутантный аллель, время существования 243—251

Неравновесие по сцеплению 315—318, 353—354

Неслучайное использование кодонов 171

Нуклеотидные последовательности, сравнительное исследование 112—120

Нуклеотидный сайт 175—176

Нутрия, инсулин 138, 139

Обратная транскриптаза 16

Обыкновенный тюлень, миоглобин 99

Однолокусный тест 304—305

Однопроходные 77

Окτανолдегидрогеназа 317

Оно закон 128

Опоссум, гемоглобин 103—186

— миоглобин 103

Орангутан 81, 136

Остракодермы (щитковые) 74

Отбор уравнивающий (балансирующий) 152, 250

— геинный 149

— генотипический 148—152, 164

— гомеостатический 36

— дарвиновский 144

— жесткий 153

— интенсивность 169

— коэффициент 62, 66, 71, 123—157, 345, 360

— массовый 28

— междемовый 28

— мягкий 153, 158

— отрицательный 70, 73, 144

— отсекающий 48, 49

— периодический 300

— по индивидуальному локусу 169—171

— ранжированный 158

— стабилизирующий 170—174

— типа *K* 77, 154

— типа *r* 77, 154

— фенотипический 145—147

— частотно-зависимый 151, 152, 158—161, 168, 289, 294—298

Оты модель 274, 276—277

— теория 141, 171, 345, 346, 362

Панселекционизм 37

Пара аллелей, квазинейтральные 69

— — псевдонейтральные 69

— — строго нейтральные 69

— — условно нейтральные 69

Пара оснований, частота мутационных замен 72, 119, 125

Параметр выживаемости 48

— популяционный 240, 285, 286

Перетасовка экзонов 360

Пиримидины 113, 115

Пирсона теория наследственности 18

Плацентарные 78

Плотность вероятности 57, 253

— — перехода 253, 254

Покрытосеменные 78, 122

Полигены жизнеспособности 155

Полиморфизм 37, 49, 152, 226, 227, 270—331, 338, 355, 356, 360

— адаптивный 32, 33

— белков 45, 284, 291, 303—357

— генетический 34—35

— на молекулярном уровне 44, 318, 341

— ферментов 45, 321

Полипептиды 40, 89, 90

Полудоминирование 62, 63, 268

Популяционная генетика 19

— — стохастические методы 22, 340

Популяция конечного размера 243—251, 335

— подразделения 333—338

— со случайным скрещиванием и перекрывающимися поколениями 54

Препроинсулин 200

Прерывистое равновесие 363

Приматы, эволюция 79—81

— филогенетическое древо 92

Принципы молекулярной эволюции 121, 127, 138

Приспособленность 143, 148, 151, 152

Проинсулин 188, 191, 342, 348

Пролин 40

Псевдогены 129, 207—211, 219—223, 332, 342—357

Птицы 76

Пуассона переменные 93

— поправка 102

— распределение 58, 98

— ряд 90, 93

Пул генов 53, 54

Пурины 113, 115

Равновесные генотипические частоты 20

Рамапитек 80

Распределение вероятности 175

— геинных частот 54, 56, 174—176

— — — стационарное 224

Райта теория смещающегося равновесия 27—29, 339

— формула 225

- Рептилии 75, 76
 Рибонуклеаза 99, 122
 — панкреатическая 95
 Рифтовая долина 81
 РНК информационная 40, 49, 110, 119, 204—218
 — транспортная 193, 194, 214—219, 349, 350
 Рыбы лососевые 133
 — чукучановые 133
- Сверхдоминирование 34, 35, 68, 150, 250, 291—298, 321, 340, 353, 355
 — маргинальное 289
 — отрицательное 151, 160, 173, 288
 — при серповидноклеточной анемии 150
 Селективная нейтральность 291—303
 — — статистические тесты 303—314
 Селективное ограничение 184—198, 350
 — преимущество 61, 62, 66, 126, 183
 Серин 40
 Сивапитек 80
 Сиетическая леталь 33
 — теория эволюции 21, 30—41, 339, 340, 341
 Скорость включения мутаций 121
 — миграции 361
 — молекулярной эволюции 97—111
 — мутирования 61, 64, 68, 123, 360
 Случайное скрещивание 149, 155, 158
 Случайный дрейф генов 27, 28, 51, 53—56, 67, 118
 — частот 241
Снига коэффициент корреляции 189, 190
 Собака, гемоглобин 88, 97, 99
 Сосудистые наземные растения 76
Спофорта модель внутрихромосомной дупликации 131, 132
 Стохастический интеграл 255
 — процесс 222—230, 252, 261
Стюарта формула 306
 Субпопуляция 60
 Сумчатые 78, 81, 85
- Такахаты модель 267
 Тахителля 83
Тейлора ряд 166
 Теория нейтральности, определение 51—52
 Терапсиды 77
 Тетраподы 75
 Тимии 40
 Тирозин 40, 187—188, 208
- Токсин змеяного яда 122
 Точковые мутации 95
 Треонин 40
 Триптофан 40
 Триптофансинтетаза А 192, 193, 214
 Тритон, гемоглобин 88, 91
 Тромбин 189, 190
 Тубулин 202, 342
 Тупайя, гемоглобин 92
- Уотсона—Крика* спаривание оснований 194, 196
 Уравнение дифференциальное 148, 232
 — диффузионное 56, 313
 — логистическое 153
 Урацил 40, 217
Уэлдона статистический метод изучения эволюции 18
- Феиаланин 40
 Фенотипический отбор 145—147, 164—171
 — — дизруптивный 145
 — — квазиротсекающий 156
 — — линейный 146
 — — направленный 145, 146, 149, 153
 — — нормализующий 146
 — — отсекающий 147, 153, 155, 156, 168
 — — стабилизирующий 145
 — — центробежный 147, 151
 — — центростремительный 146, 147
 Фермент мономерный 320, 322, 355
 — связанный с мембранами 327
 — специфичный к нескольким субстратам 323, 332, 355
 — специфичный к одному субстрату 323, 332, 355
 Фибринолтиды 94, 95, 106, 121, 122, 136, 184—190
 Фиксация гена, вероятность 39, 126, 172
 Филогенетическое древо 94, 110, 111
Фишер(a), использование стохастических методов 22
 — корреляция между родственниками 21
 — метод функциональных уравнений 22
 — модель процесса адаптации 162—164
 — теория доминантности 23—24
Фоккера—Планка диффузионное уравнение 313

Франклина—Левонтина гипотеза 315—318

ДеФриза мутационная теория 19

Функциональное ограничение 73, 184, 348, 349

— — влияние на генетическую изменчивость 320—330

Функция гипергеометрическая 56

— переходная 259

— приспособленности 145, 146, 165

— структурная дискриминантная 180

— частоты 145, 146, 165

— экспоненциальная 253

Хагедорна эффект 22

Харди—Вайнберга закон 19, 20

Холдейна—Мёллера принцип 155, 156

Хромосома, гетерозисная леталь 33

— неинвертированная 317, 354

— несущая инверсию 296

— синтетическая леталь 33

— субвитальная 33

Целакант 75

Цена естественного отбора 41, 42

Цистеин 40, 187

Цистрон 39, 65, 94, 271—273

Цитозин 39, 217

Цитохром с 94, 95, 99, 100, 106, 121, 122, 185

Цукеркандла модель, 71

Частота аллеля 353

— — описание 289—292

— генотипа 20

Человек, гемоглобин 88—99, 197

— ген α -глобина 204

— дивергенция от обезьян 80—81

— рибонуклеаза 99

— цитохром с, 99

Челюстноротые 75, 108

Черепаша, цитохром с 99, 122

Шимпанзе 136

Эволюционное расстояние 117

Эволюция, брадителя 83

— молекулярная 49

— неадаптивная 45, 46

— приспособительная 82

— синтетическая теория 21, 30—40, 339, 340, 341

— согласованная 136, 137, 352

— скорость 64, 68, 123, 185, 274, 348, 349

— — генетическая 86

— — горотелическая 42

— — морфологическая 83, 84

— — на молекулярном уровне 86—120

— — таксономическая 83

— — филогенетическая 83

— тахителы 83

— теория смещающегося равновесия 28, 29

— фенотипическая 74—86

Эйлера константа 270

Экзои 210

Экологические теории 318—320

Электроморфы 300

Электрофорез 260, 261, 263, 327, 352, 355, 357

— в полиакриламидном геле 214

— методы 39, 44, 340

Эпистаз 32, 155, 340

Эпистатическое взаимодействие генов 32, 35, 38, 354

Эстераза-С 317

Эффект «попутного транспорта» (hitch-hiking) 52, 204, 205

Эффективная численность популяции 57—61

Эффективно нейтральная мутация 67, 271—280

Юэнса—Уоттерсона однолокусный тест 304—305

Японский макак 87, 92

Оглавление

Предисловие редакторов перевода	5
Предисловие	7
Введение	9
1. От Ламарка к популяционной генетике	15
2. Развитие синтетической теории эволюции и рождение теории нейтральности	30
2.1. Становление синтетической теории и ее превращение в ортодоксальную концепцию эволюционного процесса	30
2.2. Появление мутационно-дрейфовой гипотезы (гипотезы нейтральности)	41
3. Гипотеза случайного дрейфа селективно нейтральных мутаций как эволюционная парадигма	51
3.1. Случайные факторы, действующие на селективно эквивалентные аллели	51
3.2. Случайный генетический дрейф, обусловленный ограниченной численностью популяции	53
3.3. Эффективный размер популяции	57
3.4. Нейтральные и почти нейтральные мутации	61
3.5. Популяционная динамика мутационной замены	64
3.6. О неправильном толковании теории нейтральности и ее критике	68
4. Сопоставление темпов молекулярной и фенотипической эволюции	74
4.1. Некоторые особенности фенотипической эволюции	74
4.2. Скорость эволюции на молекулярном уровне	86
5. Некоторые особенности молекулярной эволюции	121
6. Естественный отбор: определение, типы, механизмы действия	143
6.1. Значение естественного отбора	143
6.2. Фенотипический отбор по количественным признакам	145
6.3. Генотипический отбор	148
6.4. Другие виды терминологии и концепции, использующиеся при анализе естественного отбора	152
6.5. Генетический груз	154
6.6. Модель процесса адаптации, предложенная Фишером	162
6.7. Зависимость между фенотипическим и генотипическим отбором	164
6.8. Поведение мутантных аллелей в конечной популяции при стабилизирующем отборе	171
7. Структура молекул и скорость эволюции	177
7.1. Консервативная природа мутационных замен	177
7.2. Молекулярная структура и селективные ограничения	184

7.3. Синонимичные и другие «молчащие» замены	197
7.4. Быстрые эволюционные изменения псевдогенов	207
7.5. «Неслучайное» использование синонимичных кодонов	211
8. Молекулярная популяционная генетика	222
8.1. Для чего необходим стохастический подход	222
8.2. Некоторые модели аллельных мутаций	230
8.3. Время существования нейтральной мутации в популяции конечного размера	243
8.4. Интеграл произвольной функции вдоль траектории изменения частоты гена	251
8.5. Модель ступенчатого мутирования	260
8.6. Модель бесконечного числа сайтов	267
8.7. Модель эффективно нейтральных мутаций	271
8.8. Генетическое расстояние	280
9. Поддержание генетической изменчивости	284
9.1. Формулировка проблемы, факты и трактовка с позиций теории нейтральности	284
9.2. Экспериментальная проверка теории нейтральности	291
9.3. Статистические тесты селективной нейтральности	303
9.4. О некоторых альтернативных гипотезах	314
9.5. Влияние структуры молекул и функциональных ограничений на генетическую изменчивость	320
9.6. Распределение частот редких аллелей	329
9.7. Нейтральные аллели в географически подразделенной популяции	333
10. Заключение	339
Литература	365
Предметный указатель	391